

```

1   AAAATAACGA AAGAAAGGCA GAGAGGAAGT AGCGAGAGAA GAGAGAAAAT
51  GAAGTCGGCG CTGGGGGAGC CTGCAGGAGG GTGGCCAACA GTGAGGAAG
101 GTGGATTTGG CTTCTTTTCC GCACCCCGGG CGTGAAAGCC CTCTCCAACG
151 CGACCCAGAG AAATAAGTGG GTCTCGCCTG GGCAGAAAAG GAAAAGAATC
201 CAGGCGAGAG CGCGTCGCTC CTCTGTCACT GCTGCCCCCG AGGAACTCCG
251 GCTGCTTCTC ATCCCGGCCG CCTCGCGGGG CCGGACGCAG TGCCCGAGGC
301 GCCCTGCAGA TGGGGCGGGC AGGGAACGGG CGCTCCAGCT GCGGGTGACA
351 GGCGCCGGCC CGCCCGCCTG CCTGCTCAGC GCAGTGACCG GCGGGCAGA
401 GGATGCCAGG CGGAGGGACC TGGGAGCGGG ATCTGAGACT GCCGAGGCG
451 CGCTACGCTC CAACTTGCAT GGCCTAGAGA CCGCTCCAGC TCCTGGGACC
501 GCTTCAACCA GTGGAGTGAA GCTGCGCGCG GGACCTGGAG GCGGAGACCT
551 CGCAGCGCGG CTCAGAGGG GCGAGCCGGG CGCAGGAGGG GGCGCGCTTT
601 CTCCCTTCGG GTCTCAGTAA TGAGGAGACT GAGTTTGTGG TGCTGCTGA
651 GCAGGGTCTG TCTGCTGTTG CCGCCGCCCT GCGCACTGGT GCTGGCCGGG
701 GTGCCAGCTC CCTCTCGCA CCCGAGCCC TGCCAGATCC TCAAGCGCAT
751 CGGGCAGCGG GTGAGGGTGG GCGCGGTGCA CTTGAGCCC TGACCACCG
801 CCCCCCGCGC GGCCAGCCGC GCTCCGGACG ACAGCCGAGC AGGAGCCCAG
851 AGGGATGAGC CGGAGCCAGG GACTAGGCGG TCCCCGGCGC CCTCGCCGGG
901 CGCAGCCTGG TTTGGGAGCA CCCTGCATGG CCGGGGGCCG CCGGGCTCCC
951 GTAAGCCCGG GGAGGGCGCC AGGGCGGAGG CCCTGTGGCC ACGGGACGCC
1001 CTCCTATTGG CCGTGGACAA CCTGAACCGC GTGGAAGGGC TGCTACCCTA
1051 CAACCTGTCT TTGGAAGTAG TGATGGCCAT CGAGGCAGGC CTGGGCGATC
1101 TGCCACTTTT GCCCTTCTCC TCCCCTAGTT CGCCATGGAG CAGTGACCCT
1151 TTCTCCTTCC TGCAAAGTGT GTGCCATACC GTGGTGGTGC AAGGGGTGTC
1201 GGCCTGTCTC GCCTTCCCCC AGAGCCAGGG CGAAATGATG GAGCTCGACT
1251 TGGTCAGCTT AGTCTGCAC ATTCCAGTGA TCAGCATCGT GCGCCACGAG
1301 TTTCCGCGGG AGAGTCAGAA TCCCCTTCAC CTACAACCTGA GTTTAGAAAA
1351 TTCATTAAGT TCTGATGCTG ATGTCACTGT CTCAATCCTG ACCATGAACA
1401 ACTGGTACAA TTTTAGCTTG TTGCTGTGCC AGGAAGACTG GAACATCACC
1451 GACTTCCTCC TCCTTACCCA GAATAATTCC AAGTTCCACC TTGGTTCTAT
1501 CATCAACATC ACCGCTAACC TCCCCTCCAC CCAGGACCTC TTGAGCTTCC
1551 TACAGATCCA CTTGAGAGT ATTAAGAACA GCACACCCAC AGTGGTGATG
1601 TTTGCTGCGC ACATGGAAAG TATCCGGCGG ATTTTCGAAA TTACAACCCA
1651 GTTTGGGGTC ATGCCCCCTG AACTTCGTTG GGTGCTGGGA GATTCCAGA
1701 ATATGGAGGA ACTGAGGACA GAGGGTCTGC CCTAGGACT CATGCTCAT
1751 GGAAAAACAA CACAGTCTGT CTTTGAGCAC TACGTACAAG ATGCTATGGA
1801 GCTGGTCGCA AGAGCTGTAG CCACAGCCAC CATGATCCAA CCAGAACTTG
1851 CTCTCATTCC CAGCAGGATG AACTGCATGG AGGTGGAAC TACAAATCTC
1901 ACTTCAGGAC AATATTTATC AAGGTTTCTA GCCAATACCA CTTTCAGAGG
1951 CCTCAGTGGT TCCATCAGAG TAAAAGGTTT CACCATCGTC AGCTCAGAAA
2001 ACAACTTTTT CATCTGGAAT CTTCAACATG ACCCATAGGG AAAGCCAATG
2051 TGGACCCGCT TGGGCAGCTG GCAGGGGAGA AAGATTGTCA TGGACTATGG
2101 AATATGGCCA GAGCAGGCCC AGAGACACAA AACCCACTTC CAACATCCAA
2151 GTAAGCTACA CTTGAGAGTG GTTACCCTGA TTGAGCATCC TTTTGTCTTC
2201 ACAAGGGAGG TAGATGATGA AGGCTGTGTC CCTGCTGGCC AACTCTGTCT
2251 AGACCCCATG ACTAATGACT CTTCCACACT GGACAGCCTT TTAGCAGCC
2301 TCCATAGCAG TAATGATACA GTGCCCATTA AATTCAAGAA GTGCTGCTAT
2351 GGATATTGCA TTGATCTGCT GGAAAAGATA GCAGAAGACA TGAACTTTGA
2401 CTTGACCTC TATATTGTAG GGGATGGAAA GTATGGAGCC TGGAAAAATG
2451 GGCCTGGAC TGGGCTAGTG GGTGATCTCC TGAGAGGGAC TGCCCATG
2501 GCAGTCACTT CCTTTAGCAT CAATACTGCA CGGAGCCAGG TGATAGATT
2551 CACCAGCCCT TTCTTCTCCA CCAGCTTGGG CATCTTAGTG AGGACCCGAG
2601 ATACAGCAGC TCCCATTGGA GCCTTCATGT GGCCACTCCA CTGGACAATG
2651 TGGCTGGGGA TTTTGTGGC TCTGCACATC ACTGCCGTCT TCCTCACTCT
2701 GTATGAATGG AAGAGTCCAT TTGGTTTGAC TCCAAGGGG CGAAATAGAA
2751 GTAAAGTCTT CTCCTTTTCT TCAGCCTTGA ACATCTGTGA TGCCCTCTTG
2801 TTTGGCAGAA CAGTGGCCAT CAAACCTCCA AAATGTTGGA CTGGAAGGTT
2851 TCTAATGAAC CTTTGGGCCA TTTTCTGTAT GTTTTGCTT TCCACATACA
2901 CGGCAAACTT GGCTGCTGTC ATGGTAGGTG AGAAGATCTA TGAAGAGCTT
2951 TCTGGAATAC ATGACCCCAA GTTACATCAT CCTTCCCAAG GATTCCGCTT
3001 TGGAAGTGTG CGAGAAAGCA GTGCTGAAGA TTATGTGAGA CAAAGTTTCC
3051 CAGAGATGCA TGAATATATG AGAAGGTACA ATGTTCCAGC CACCCCTGAT
3101 GGAGTGGAGT ATCTGAAGAA CAATCCAGAG AACTAGACG CCTTCATCAT

```

FIGURE 1, page 1 of 4

3151	GGACAAAGCC	CTTCTGGATT	ATGAAGTGTC	AATAGATGCT	GA CTGCAAAC
3201	TTCTCACTGT	GGGGAAGCCA	TTTGCCATAG	AAGGATACGG	CATTGGCCCTC
3251	CCACCCAACT	CTCCATTGAC	CGCCAACATA	TCCGAGCTAA	TCAGTCAATA
3301	CAAGTCACAT	GGGTTTATGG	ATATGCTCCA	TGACAAGTGG	TACAGGGTGG
3351	TTCCCTGTGG	CAAGAGAAGT	TTTGCTGTCA	CGGAGACTTT	GCAAATGGGC
3401	ATCAAACACT	TCTCTGGGCT	CTTTGTGCTG	CTGTGCATTG	GATTTGGTCT
3451	GTCCATTTTG	ACCACCATTG	GTGAGCACAT	AGTATACAGG	CTGCTGCTAC
3501	CACGAATCAA	AAACAAATCC	AAGCTGCAAT	ACTGGCTCCA	CACCAGCCAG
3551	AGATTACACA	GAGCAATAAA	TACATCATTT	ATAGAGGAAA	AGCAGCAGCA
3601	TTTCAAGACC	AAACGTGTGG	AAAAGAGGTC	TAATGTGGGA	CCCCGTCAGC
3651	TTACCTGATG	GAATACTTCC	AATCTGAGTC	ATGACAACCG	ACGGAAATAC
3701	ATCTTTAGTG	ATGAGGAAGG	ACAAAACCAG	CTGGGCATCC	GGATCCACCA
3751	GGACATCCCC	CTCCCTCCAA	GGAGAAGAGA	GCTCCCTGCC	TTGCGGACCA
3801	CCAATGGGAA	AGCAGACTCC	CTAAATGTAT	CTCGGAACCTC	AGTGATGCAG
3851	GAACTCTCAG	AGCTCGAGAA	GCAGATTCAG	GTGATCCGTC	AGGAGCTGCA
3901	GCTGGCTGTG	AGCAGGAAAA	CGGAGCTGGA	GGAGTATCAA	AGGACAAGTC
3951	GGACTTGTGA	GTCCTAGGTG	ACCACACTGC	TTCCCTTTCT	CAGTTCTCTGA
4001	CCTTCCCTCTG	AGCCCTTGAG	ACACTTTGTA	ATGCTCTTTT	GTA ACTATCG
4051	ACAAAGGTGT	GGGGAAGCTG	AGGTCTAGGT	CTTCTTAAAG	GTCAAGTCTG
4101	CTCTCCCTCG	CCTAAAGTGC	AGCAGCAGCT	CCTCTCAAGC	TCACTCTCTA
4151	GGTCTCCAGG	G TAGGAGTGT	TTTTCTAGCA	AGAATCTTAG	TCAGGAGTAA
4201	GCTCTGTGCG	AGAGATCTGT	GAATAACCAG	ATAACCCCAG	CTGCCGTTAA
4251	CCTTTTCAAC	AGGTGCCACA	GTAATATTTT	TGGTTTTTAG	CCCTTTCTCT
4301	GCACTACCAA	CAAGAGATAA	AATTGTTACT	CACACTTATG	TCTTACTGGG
4351	TTGCTGGTTT	TCATCGTAAC	ACAGAACGAG	GTTATCTAGG	GTTGTAGCTT
4401	TTGATACAAC	TCCCCGATCT	AGATTATATC	CTACATTCTG	AATGGGGAGC
4451	AGGTAAGAGC	AGAGCACCTC	CCACTGGGGG	TGGGGTATTT	AAAAATTAAC
4501	TCATTAGTAT	CATAAACGTC	AAGGATTGAT	TGGACCAGGC	AAGAGCCATG
4551	TTTTTTAGAA	GGTTCTGGAT	CTCTGACTCC	ATCCTGACTG	TTTAGTAAGA
4601	GCATGCTTAC	ACCTACTGTG	GAAAAGGGGA	GGGGATGTGG	TAAGCGGAAA
4651	CAGAAGACAG	GCAGCAGAGG	CATTAAAAAT	GCATACCATG	CTTTCAGAAC
4701	AAAAGCTCTG	GGCCAGAAAG	GCAATTGTGG	TAAAAAATGA	ATAAGACTAC
4751	TTCTAATGTA	ACTAAGCATC	TCCACTATGG	TGTGTGCCTT	TTATAAAGGA
4801	AAAGAGAGAA	AAAGGCAAAG	CAAGGTTGTG	GCCTTAGGTT	GGACCTGGAA
4851	TATCCCTTAT	TGCCTATAAT	GGAATATGTG	ACACTGTGGG	TGAAATGTTT
4901	TACACACCAC	ACACTAGGCC	ATTTTCAGAT	CAGCAGTCAC	CCATCGCTTA
4951	GCATAGAAAT	CCCAAAACCT	CCAGCCCGGG	AACACTATAA	GCTTCGACCA
5001	TTCAGGAATC	TGCCCTGCAC	TTTGCAATATC	TGTATAGAAA	ATCAAGTCAA
5051	TCCCCCATCC	TCACACCCAC	TCATCTCTGA	GGAGCTATGA	ACTGGTTTTG
5101	GTCCCTCTAA	TGATCCTCCA	GCCTCATCTA	ATGCCCCCCA	AAGACTGATA
5151	CAAGTAACCT	CCCCTCTGCT	TAGGTGTCAC	TTTCTCAGCA	TATCAAGTTT
5201	AGGCAGCAAG	GGAAAGGAAT	ATGGGTCAGT	TCTCAAATGT	CAATGTAGAT
5251	AAGAGTCATC	TAGTAGAGAA	CTCATCAGAG	TGCGGATTGC	CAAGACCCCT
5301	CTCCAGAGAT	TATGGGGTTG	GGGGTGGAGG	TCTAGAGGTG	AGCTCAGAAA
5351	CCTACTGTTA	ACCAACACCC	CCAAGTGACT	GACACAGGTG	GTCTAAAAAT
5401	TACTTTTCTA	GAAACACCAT	TCTGGAAGTT	TGGCTGCCCA	CAGGCAGGAG
5451	GAGAAGCATG	AAGAGAAAAC	CTGTTTGAGA	AGTTTTGTTT	TGTTTTGTTT
5501	TGCTTTTTTAA	TAATTTTAGC	ACACATCTGC	TGACTCTCCT	TCAACATCCT
5551	CACCCCCACC	CCTGGGCACC	ATTTAGGACA	AGACTTCCTT	ATTTATCAAT
5601	TACTTGATTT	ATCTTCTCAG	GACTCATTGT	TCCACCCCCA	ACCAATTTGA
5651	ATGCCCTACAA	TAAGTTCAGG	AGCTGTGCCA	AGCACTTTCC	TCTTTTACAG
5701	CTGGAGATCA	CTGGAAAGGT	GTCTCAGTCA	CAAAACTTCT	CCCTCTACTA
5751	CTGGATGAAA	TGTCTGCATT	TCCACCAAAA	TCTACCCAGT	CACCCAGGGA
5801	ATAACAACCT	AAGCTGTAGT	TAGATAACAC	CTAGTGATTA	ATTGGCTGAG
5851	AAAACCCTGG	AGTGGAGGGA	GGCTCAGAGA	TACTGATATG	GATGTGGGAG
5901	GGCTCTAAAG	TTAGAGGTCA	CCAACCTCCAC	AGATGAAACA	GTTCAATAAT
5951	GAGGAAACAG	GTGAGCCCTG	AAAACACAAA	AGGACAGTTC	TGTGTTGAAA
6001	CACCCCATCC	CCTCACGTTT	TCACCCAGG	CCCAGAAGTA	GTTTGCAACT
6051	GCCTTTGGAA	GATTTTGCCC	CTTAGCCATC	CCCACCCACT	TGTACCAGCT
6101	AAGAATGCTG	GAGACTCTGC	CACCATGCTC	TGCGTGCCCC	TGAACCTCTG
6151	TGCAGCCCGG	AAGGCTGATG	TACAGGTGTA	CCTCAATCCA	CATTACAGCC
6201	ATGCTCCTAA	TGTACATGGA	CATTTTGTGA	ACTCAGCTCA	TATTCTGACT
6251	GTATTTTGAGA	AGCTGGCTGT	TTAAGGGAAC	CCAGAAGTGA	ATTCTTTTGT

FIGURE 1, page 2 of 4

```

6301 AAAGTAAAGC ACCCTTTTGT AATGCAATTA ATTATCCCTT AATGTATCTG
6351 TTTTGTAAAGT CTGCATTTT GTATATCGGA TTTACCTTAA GCTTCTCTAG
6401 TGAGGCATTC TGAGCAGTGG TGATCACATG CCAGATCGCC CTGCCTATCC
6451 ACAAAGTAGA TGACCAATGC ACGCTCCTCA AACATCTTTG GAGGAACTAC
6501 CTGGCCAAAA CACTGGCCAG GATGCAGCAA GCAGCAGCAG GGGCTGACAG
6551 CAGGCTTACT GCCATCAACA TTGCTTGAAA TGCCTCTATG TTCTGAATAA
6601 AGAAAAACCA TAATTGCTTG TGGTGAAACG AAGCAGTCTT CATGTTAAGT
6651 AGCAATGGTT ATTTTATTG GTAGTAACTG AACAGTGTGTT TGCAATTTGT
6701 GAAACAGTGT ATTGTGTTT GTAAAAATGAT GTCATGAAAT GGTGGGTCCT
6751 TGGAAACCTC CTTTCCGTTT AGCTCTGCCT CTGTTCTTTC AACTCCTTTG
6801 AGGCTCAAAA AAAACACAAA GATCAGAAGC CTTCAGATAG AGGGTGGTAT
6851 TCTGGTAAAG AAGAAAGAGA TAAGGGACGC TACCTTGCTT TTCTGGCACA
6901 GGAAGCACAT GATAAAGCAT GCTCAGATGA GCTGGAACAG ATATAGCTAC
6951 CTGGTTCGTG TAAATAAGAA TAATCAAGGC CCCAGAGTGT GTATGCTTCC
7001 AGGTGGAGGA GAAAGGGGAA TCTCCCAAAA TTTAAAAACA AATTGGAAGA
7051 ATAACCAGGA CAGCCAAGTG AAGCAGCCAC AGGGACCCAA GCAGTCGAGG
7101 TCTTTAATGT GCCTGGAGAT GACTCTCTGC TATTCATGAA TCTTGCTATT
7151 GCACAAACCC TATCAAGAGC TGCTGCTTCC CTTCCAGCCA GAAAAGTGGT
7201 AAGCGGAGCA AGTGCCAAGC AGAACAGACC TTATCATCTG GGTAACAGAC
7251 TTCTCAGTGT TGGTGCTGTG TCTGTTAGAG CCTTAGAGCA AGTTAAGCAC
7301 TTCCTTGGTG TGGGTAAAGA ATAAAGGGGA AAGAACTAC TTTAGAGCCT
7351 CTTTTTCTCC CAACTCATAT TTTTGATAGG AAAACAGAA AACCCATCCA
7401 GTTCTTCAGA AATTGCTTTC TAGGCATTAA TACTACTTTA CTATCTATAC
7451 TGTTTAGTTA TTCCTTTCTT TACCCACCTA AACTATCCAT CTAATCCAGG
7501 ATTCCTCAC TCTTTTTTTT TAGTTACTAA TCATTTTATG AAAATAATGT
7551 ATTTATAAGT ATTTTCTTAA GGTTTGTGAA GAGTATTTGC ATTGTGTCTT
7601 CATTTTAATG TGTTTGCAAT CGCTCCGCTC CAGGAAGAAC GGAAATGCTG
7651 TCTTGTGAGC ATGAAGTGAA CGGGCTGTTT TGCTCCAGCC ACTTTTCTTG
7701 TACAACCACA TGGATGGATT AGATGTCTC AGGTCTTTTC CATCTTCAGT
7751 TTCTATGACT GTGGAATAAA TGTTCAAGATA GAAACTTCA (SEQ ID NO:1)

```

FEATURES:

5'UTR: 1-619
 Start: 620
 Stop: 3965
 3' UTR: 3968-7789

HOMOLOGOUS PROTEIN:

Top 10 BLAST Hits:

gi 7514020 pir	T31068 N-methyl-D-aspartate receptor homolog NM...	1948	0.0
gi 5305435 gb	AAD41650.1 AF073379_1 (AF073379) N-methyl-D-aspar...	1936	0.0
gi 3025446 gb	AAC12680.1 (AC004528) R32184_2 [Homo sapiens]	906	0.0
gi 3822016 gb	AAD11811.1 (AF061945) NMDA receptor-like long va...	456	e-127
gi 286234 dbj	BAA02498.1 (D13211) N-methyl-D-aspartate recepto...	304	2e-81
gi 2155310 gb	AAB58801.1 (AF001423) N-methyl-D-aspartate recep...	302	1e-80
gi 6980982 ref	NP_036705.1 glutamate receptor, ionotropic, N-...	302	2e-80
gi 6680097 ref	NP_032196.1 glutamate receptor, ionotropic, NM...	300	4e-80
gi 6680099 ref	NP_032197.1 glutamate receptor, ionotropic, NM...	299	8e-80
gi 4099613 gb	AAD00659.1 (U88963) N-methyl-D-aspartate recepto...	299	8e-80
gi 228950 prf	1814459A D-MeAsp receptor:SUBUNIT=epsilon2 [Mus ...	299	8e-80
gi 548372 sp	Q00960 NME2_RAT GLUTAMATE [NMDA] RECEPTOR SUBUNIT ...	299	8e-80
gi 6980984 ref	NP_036706.1 glutamate receptor, ionotropic, N-...	299	8e-80

blast to dbEST:

ESTs (from GenBank EST division)

AI267842 aq35f07.x1 Stanley Frontal SN pool 1 Homo sapiens cDNA clone IMAGE:2032933, mRNA
sequence [Homo sapiens]

EXPRESSION INFORMATION FOR MODULATORY USE:

Expression information from BLAST EST hit:

AI267842 228 bp mRNA EST 17-NOV-1998 frontal lobe of the brain

Expression information from cDNA library screening:

Human Liver

Human Brain

Human Fetal Brain

Human Bone Marrow

Human Adrenal Gland

Human Heart

Human Mammary Gland

Human Pituitary

Human Testis

```

1  MRRLSLWLL SRVCLLLPPP CALVLAGVPS SSSHPQPCQI LKRIGHAVRV
51  GAVHLQPWTT APRAASRAPD DSRAGAQRDE PEPGTRRSPA PSPGARWLGS
101 TLHGRGPPGS RKPGEARAE ALWPRDALLF AVDNLNRVEG LLPYNLSLEV
151 VMAIEAGLGD LPLLPFSSPS SPWSSDPFSF LQSVCHTVVV QGVSALLAFP
201 QSQGEMMELD LVSLVLHIPV ISIVRHEFPR ESQNPLHLQL SLENSLSSDA
251 DVTVSILTMTN NWYNFSLLLC QEDWNITDFL LLTQNNSKFH LGSIIINITAN
301 LPSTQDLLSF LQIQLESIKN STPTVVMFGC DMESIRRIFE ITTQFGVMPP
351 ELRWVLGDSQ NMEELRTEGL PLGLIAHGKT TQSVFEHYVQ DAMELVARAV
401 ATATMIQPEL ALIPSTMNCM EVETTNTLSTG QYLSRFLANT TFRGLSGSIR
451 VKGSTIVSSE NNFFIWNLQH DPMGKPMWTR LGSWQGRKIV MDYGIWPEQA
501 QRHKTHFQHP SKLHLRVVTL IEHPFVFTRE VDDEGLCPAG QLCLDPMTND
551 SSTLDSLFSS LHSSNDTVPI KFKKCCYGYC IDLLEKIAED MNFDFDLYIV
601 GDGKYGAWKN GHWTGLVGDL LRGTAHMAVT SFSINTARSQ VIDFTSPFFS
651 TSLGILVRTR DTAAPIGAFM WPLHWTMWLG IFVALHITAV FLTLYEWKSP
701 FGLTPKGRNR SKVFSFSSAL NICYALLFGR TVAIKPPKCW TGRFLMNLWA
751 IFCMFCLSTY TANLAAMVG EKIYEELSGI HDPKLHHPQS GFRFGTVRES
801 SAEDYVRQSF PEMHEYMRRY NVPATPDGVE YLKNDEPEKLD AFIMDKALLD
851 YEVSIDADCK LLTVGKPFAP EGYGIGLPPN SPLTANISEL ISQYKSHGFM
901 DMLHDKWYRV VPCGKRFAV TETLQMGIKH FSGLFVLLCI GFGLSILTTI
951 GEHIVYRLLL PRIKNKSKLQ YWLHTSQRLLH RAINTSFIEE KQQHFKTKRV
1001 EKRSNVGPRQ LTVWNTSNLS HDNRRKYIFS DEEGQNQLGI RIHQDIPLPP
1051 RRRELPAVRT TNGKADSLNV SRNSVMQELS -ELEKQIQVIR QELQLAVSRK
1101 TELEEQRTS RTCES (SEQ ID NO:2)

```

FEATURES:

Functional domains and key regions:

PDOC00001 PS00001 ASN_GLYCOSYLATIONN-glycosylation site

Number of matches: 16

```

1 145-148 NLSL
2 264-267 NFSL
3 275-278 NITD
4 285-288 NNSK
5 296-299 NITA
6 426-429 NLTS
7 439-442 NTTF
8 549-552 NDSS
9 565-568 NDTV
10 709-712 NRSK
11 886-889 NISE
12 965-968 NKSK
13 984-987 NTSF
14 1015-1018 NTSN
15 1018-1021 NLSH
16 1069-1072 NVSR

```

[2]

PDOC00004 PS00004 CAMP_PHOSPHO_SITEcAMP- and cGMP-dependent protein kinase phosphorylation site

2-5 RRLS

[3]

PDOC00005 PS00005 PKC_PHOSPHO_SITEProtein kinase C phosphorylation site

Number of matches: 14

```

1 85-87 TRR
2 110-112 SRK
3 1098-1100 SRK
4 317-319 SIK
5 334-336 SIR
6 448-450 SIR
7 441-443 TFR
8 334-336 SIR
9 448-450 SIR
10 636-638 TAR
11 704-706 TPK
12 741-743 TGR

```

```

13      796-798 TVR
14      976-978 SQR
-----[4]
PDOC00006 PS00006 CK2_PHOSPHO_SITECasein kinase II phosphorylation site
Number of matches: 17
  1      202-205 SQGE
  2      248-251 SDAD
  3      303-306 STQD
  4      383-386 SVFE
  5      519-522 TLIE
  6      552-555 STLD
  7      563-566 SSND
  8      693-696 TLYE
  9      796-799 TVRE
 10      800-803 SSAE
 11      801-804 SAED
 12      809-812 SFPE
 13      949-952 TIGE
 14      986-989 SFIE
 15     1030-1033 SDEE
 16     1080-1083 SELE
 17     1101-1104 TELE
-----[5]
PDOC00007 PS00007 TYR_PHOSPHO_SITETyrosine kinase phosphorylation site
      1099-1106 RKTELEY
-----[6]
PDOC00008 PS00008 MYRISTYLN-myristoylation site
Number of matches: 10
  1      27-32 GVPSSS
  2      292-297 GSIINI
  3      369-374 GLPLGL
  4      430-435 GQYLSR
  5      444-449 GLSGSI
  6      482-487 GSWQGR
  7      535-540 GLCPAG
  8      606-611 GAWKNG
  9      680-685 GIFVAL
 10      876-881 GLPPNS
-----[7]
PDOC00009 PS00009 AMIDATIONAmidation site
Number of matches: 2
  1      485-488 QGRK
  2      913-916 CGKR
-----[8]
PDOC00017 PS00017 ATP_GTP_AATP/GTP-binding site motif A (P-loop)
      373-380 GLIAHGKT

```

Membrane spanning structure and domains:

1	13	33	1.496	Certain
2	182	202	1.149	Certain
3	206	226	0.679	Putative
4	251	271	0.628	Putative
5	399	419	0.694	Putative
6	639	659	0.949	Putative
7	675	695	2.050	Certain
8	711	731	1.052	Certain
9	744	764	1.803	Certain
10	931	951	2.294	Certain

BLAST Alignment to Top Hit:

```
>gi|7514020|pir||T31068 N-methyl-D-aspartate receptor homolog NMDAR-L
- rat >gi|1050330|gb|AAA99501.1| (L34938) ionotropic
glutamate receptor [Rattus norvegicus]
>gi|2160125|gb|AAB58957.1| (U29873) NMDAR-L [Rattus
norvegicus] Length = 1115
Score = 1948 bits (4990), Expect = 0.0
Identities = 941/1089 (86%), Positives = 992/1089 (90%)
Query: 27 GVPSSSSHPQPCQILKRIGHAVRVGAVHLQPWTTXXXXXXXXXXXXXXXXXGAQRDEPEPGTR 86
GVPSSSSHPQPCQILKRIGHAVRVGAVHLQPWTT GAQRD+PE GT
Sbjct: 27 GVPSSSSHPQPCQILKRIGHAVRVGAVHLQPWTTAPRAASRAQEGGRAGAQRDDPESGTW 86

Query: 87 RSPAPSPGARWFGSTLHGRGPPGSRKPGEGARAEALWPRDALLFAVDNLNRVEGLLPYNL 146
R PAPS GARW GS LHGRGPPGSRK GEGA AE LWPRDALLFAV+NLNRVEGLLPYNL
Sbjct: 87 RPPAPSQGARWLGSALHGRGPPGSRKLGEGAGAETLWPRDALLFAVENLNRVEGLLPYNL 146

Query: 147 SLEVVMIEAXXXXXXXXXXXXXXXXXXXXXXXXXXQSVCHTVVVQGVSAALLAFPQSQGEM 206
SLEVVMIEA QSVCHTVVVQGVSAALLAFPQSQGEM
Sbjct: 147 SLEVVMIEAGLDLPLMPFSSPSSPWSSDPFSFLQSVCHTVVVQGVSAALLAFPQSQGEM 206

Query: 207 MELDLVSLVLHIPVISIVRHEFPRESQNPLHLQLSLENSLSSDADVTVSILTMNNWYNFS 266
MELDLVS VLHIPV+SIVRHEFPRESQNPLHLQLSLENSLSSDADVTVSILTMNNWYNFS
Sbjct: 207 MELDLVSSVLHIPVLSIVRHEFPRESQNPLHLQLSLENSLSSDADVTVSILTMNNWYNFS 266

Query: 267 LLLCQEDWNITDFLLLTQNNSKFHLGSIINITANLPSTQDLLSFLQIQLESIKNSTPTVV 326
LLCQEDWNITDFLLLT+NNSKFHL S+INITANL ST+DLLSFLQ+Q+++I+NSTPT+V
Sbjct: 267 LLLCQEDWNITDFLLLTENNSKFHLESVINITANLSSTKDLLSFLQVQMDNIRNSTPTMV 326

Query: 327 MFGCDMESIRRIFEITTQFGVMPPELRWVLGDSQNMEELRTEGLPLGLIAHGKTTQSVFE 386
MFGCDM+SIR+IFE++TQFG+ PPEL WVLGDSQN+EELRTEGLPLGLIAHGKTTQSVFE
Sbjct: 327 MFGCDMSIRQIFEMSTQFGLSPPELHWVLGDSQNVEELRTEGLPLGLIAHGKTTQSVFE 386

Query: 387 HYVQDAMELVARAVATATMIQPELALIPSTMNCEVETTNTLSGQYLSRFLANTTFRGLS 446
+YVQDAMELVARAVATATMIQPELAL+PSTMNCM+V+TTNLTSGQYLSRFLANTTFRGLS
Sbjct: 387 YVQDAMELVARAVATATMIQPELALLPSTMNCMDVKTTNLTSGQYLSRFLANTTFRGLS 446

Query: 447 GSIRVKGSTIVSSENNFFIWNLQHDPMGKPMWTRLGSWQGRKIVMDYGIWPEQAQRHKTH 506
GSI+VKGSTI+SSENNFFIWNLQHDPMGKPMWTRLGSWQG +IVMD GIWPEQAQRHKTH
Sbjct: 447 GSIKVKGSTIISSENNFFIWNLQHDPMGKPMWTRLGSWQGGRIVMDSGIWPEQAQRHKTH 506

Query: 507 FQHPSKLHLRVVTLIEHPFVFTREVDDDEGLCPAGQLCLDPMXXXXXXXXXXXXXXXXXXXX 566
FQHP+KLHLRVVTLIEHPFVFTREVDDDEGLCPAGQLCLDPM
Sbjct: 507 FQHPNKLHLRVVTLIEHPFVFTREVDDDEGLCPAGQLCLDPMTNDSSMLDRLFSSLHSSND 566

Query: 567 XVPIKFKKCCYGYCIDLLEKIAEDMNFDFDLYIVGDGKYGAWKNGHWTGLVGDLLRGTAH 626
VPIKFKKCCYGYCIDLLE++AEDMNFDFDLYIVGDGKYGAWKNGHWTGLVGDLL GTA+
Sbjct: 567 TVPIKFKKCCYGYCIDLLEQLAEDMNFDFDLYIVGDGKYGAWKNGHWTGLVGDLLSGTAN 626

Query: 627 MAVTSFSINTARSQVIDFTSPFFSTSLGILVRTRDTAAPIGAFMWPLHWTMWLGIFVALH 686
MAVTSFSINTARSQVIDFTSPFFSTSLGILVRTRDTAAPIGAFMWPLHWTMWLGIFVALH
Sbjct: 627 MAVTSFSINTARSQVIDFTSPFFSTSLGILVRTRDTAAPIGAFMWPLHWTMWLGIFVALH 686

Query: 687 ITAVFLTLIEWKSPFGLTPKGRNRSKVFSFSSALNICYALLFGRTVAIKPPKCWTGRFLM 746
ITA+FLTLIEWKSPFG+TPKGRNR+KVFSFSSALN+CYALLFGRT AIKPPKCWTGRFLM
Sbjct: 687 ITAIFLTLIEWKSPFGMTPKGRNRNRSKVFSFSSALNVCYALLFGRTAAIKPPKCWTGRFLM 746

Query: 747 NLWAIFCMFCLSTYTANLAAMVGEKIYEELSGIHDPKLHHP SQGFRFGTVRESSAEDYV 806
NLWAIFCMFCLSTYTANLAAMVGEKIYEELSGIHDPKLHHP SQGFRFGTVRESSAEDYV
Sbjct: 747 NLWAIFCMFCLSTYTANLAAMVGEKIYEELSGIHDPKLHHP SQGFRFGTVRESSAEDYV 806

Query: 807 RQSFPEMHEYMRRYNVPATPDGVEYLKNNPEKLDAFIMDKALLDYEVSIDADCKLLTVGK 866
RQSFPEMHEYMRRYNVPATPDGV+YLKN+PEKLDAFIMDKALLDYEVSIDADCKLLTVGK
Sbjct: 807 RQSFPEMHEYMRRYNVPATPDGVQYLKNDPEKLDAFIMDKALLDYEVSIDADCKLLTVGK 866
```

Query: 867 PFAIEGYGIGLPPNSPLTANISELISQYKSHGFM DMLHDKWYRVVPCGKR SFAVTETLQM 926
PFAIEGYGIGLPPNSPLT+NISELISQYKSHGFM D+LHDKWY+VVP CGKR SFAVTETLQM
Sbjct: 867 PFAIEGYGIGLPPNSPLTSNISELISQYKSHGFM DVLHDKWYKVVP CGKR SFAVTETLQM 926

Query: 927 GIKHFSGLFVLLCIGFGLSILTTIGEHIVRLLLPR IKNKSKLQYWLHTSQR LHRAINTS 986
GIKFHSGLFVLLCIGFGLSILTTIGEHIV+RLLLPR IKNKSKLQYWLHTSQR HRA+NTS
Sbjct: 927 GIKHFSGLFVLLCIGFGLSILTTIGEHIVRLLLPR IKNKSKLQYWLHTSQR FHRAINTS 986

Query: 987 FIEEKQQHF KTKRVEKRSNVGPRQLTVWNTSNLSHDNR RKYIFSDEEGQNQLGIRIHQDI 1046
F+EEKQ KTKRVEKRSN+GP+QL VWNTSNLSHDN+RKYIF+DEEGQNQLG + HQDI
Sbjct: 987 FVEEKQPRSKTKRVEKRSNLGPQQLMVWNTSNLSHDNR KYIFNDEEGQNQLGTQAHQDI 1046

Query: 1047 XXXXXXXXXXXXXTTNGKADSLNVS RNSVMQELSELEKQIQVIRQELQLAVSRKTELEEY 1106 (residues 27-
1106 of SEQ ID NO:2)

TTNGKADSLNV+R+SV+QELSELEKQIQVIRQELQLAVSRKTELEEY

Sbjct: 1047 PLPQRRREL PASLT TNGKADSLNVTRSSVIQELSELEKQIQVIRQELQLAVSRKTELEEY 1106 (SEQ ID NO:4)

Hmmer search results (Pfam):

Model	Seq-from	Seq-to	HMM-from	HMM-to	Score	E-value	Description
lig_chan	674	952	1	304	199.6	1.7e-57	Ligand-gated ion channel

1 ATGGAACTT TAGCTCATGA ATCAAAACAA CCCTCCAGAG CTAAGGCCA
 51 GCTGTATTG CATAACAATT TAGCAGATCC AAACAGCAGG GCAAGGTCGG
 101 GTGAAATAAG TTGCCAAGGT CATGGTCATG AAGTAGTATT AGACTCAGAA
 151 AGGCTGATCC CCAGTGCTTG CTCCACCCCA TGGATCTCTC CTACCCTCCT
 201 TCTAAACGAT ACTGTGGGAT AAAATAAAAT TAATCTACTG TATATGTGCA
 251 AACCACAGGC CTGCCCTTAA CTCTTCTCTT ACCTTCTAGT TTCAGATTAT
 301 TCAAATCATG GAGGAAAAGA TTAGATCACA ACACGTTGAC TTCACTGTAT
 351 TACCATACAA ATGAAATAAC TTAGTACAAA CTGTGATCTG GGGACTCTTG
 401 ATCTAACTG GGAAGTGTG TTGACTGCAT TTTAACTCT AAAAGTATTT
 451 TGAACTCTT TAATTTCTTG AACTGAAAAA ATTGCTTTGA ATTCACTTTG
 501 TTTTAACTCT GAGAACCTAA AAACAGGGAT TCTTTAAAAA AAAAAATGCA
 551 AAGGCTCACA TCAGCAACAG AAAGAAGCTG AGGAGATAAA AATGTGTAAA
 601 TAATTTCTTAC TTTAATACCC TTAGCTAGAA AAACCTTAAA AGCGACACAT
 651 CCAGAAGCTC GTTAAGTCAC AGCCTCTTTG AACCTATTTT AGTGAACCAC
 701 CGAATTTTCA ATCCCTCAGG TGCAGCTCTG AATTCAGAAT TCTACCGGC
 751 TCATAGTCCT ATTTTCTTTC TTAGGTTTTA GGGAAATTTG CAAACTATGA
 801 CGCCAGCCT TTGAGGGGAG AGGACTTTCC AGGGGCGCGG GATGTGCCAC
 851 TCGGGAATCT CACCAACAGT GGGCGTTTAG CGCAGCCAAG CGACAGGCAG
 901 GCGCCAGGCG TCAGCAACAG GGAGGCGCGG GCTGAGGCGG GGAGAACTTT
 951 GGCGCTCGGA GCAGAGCCAC CCTTTGCTGG CCAGTCGCGT TGCTCTCCG
 1001 AGGAAGCAAG CGGCGGTGGC GACTCGGTGG AAAATAACG AAAGAAAGGC
 1051 AGAGAGGAAG TAGCGAGAGA AGAGAGAAAA TGAAGTCGGC GCTGGGGGAG
 1101 CCTGCAGGAG GGTGGCCAAC AGTGGAGGAA GGTGGATTTG GCTTCTTTTC
 1151 CGCACCCCGG GCGTGAAAGC CCTCTCCAAC GCGACCCAG GAAATAAGTG
 1201 GGTCTCGCCT GGGCAGAAAA GGAAAGAAT CCAGGCGAGA GCGGCTCGCT
 1251 CCTCTGTAC TGCTGCCCGC GAGGAAGTCC GGCTGCTTCT CATCCCGGCC
 1301 GCCTCGCGGG GCGGACGCA GTGCGGAGG CGCCTGCGG ATGGGGCGGG
 1351 CAGGGAACGG GCGCTCCAGC TGCGGGTGAC AGGCGCCGGC CCGCCGCT
 1401 GCCTGTCTAG CGCAGTGACC GGGCGGGCAG AGGATGCCAG GCGAGGGAC
 1451 CTGGGAGCGG GATCTGAGAC TGCCGAGGC GCGTACGCT CCAACTTGCA
 1501 TGGCCTAGAG ACCGCTCCAG CTCTGGGAC CGCTTACCG AGTGGAGTGA
 1551 AGCTGCGCGC GTGACCTGGA GGCGGAGACC TCAGGACGCG GCTGCAGAGG
 1601 GGCGAGCCGG GCGCAGGAGG GGGCGCGCTT TCTCCCTGCG GGTCTCAGTA
 1651 ATGAGGAGAC TGAGTTTGTG GTGGCTGCTG AGCAGGGTCT GTCTGCTGTT
 1701 GCGCCCGCCC TGCGCACTGG TGCTGGCCGG GGTGCCAGC TCCTCTCGC
 1751 ACCCGCAGCC CTGCCAGATC CTCAAGCGCA TCGGCGACGC GGTGAGGGTG
 1801 GGCGCGGTGC ACTTGACGCC CTGGACCACC GCGCCCGCG CCGCCAGCCG
 1851 CGCTCCGAG GACAGCCGAG CAGGAGCCCA GAGGATGAG CCGGAGCCAG
 1901 GGAGTAGGCG GTCCCGGCG CCCTCGCCGG GCGCACGCTG GTTGGGGAGC
 1951 ACCCTGCATG GCGGGGGGCC GCGGGGCTCC CGTAAGCCCG GGGAGGGCGC
 2001 CAGGGCGGAG GCGCTGTGGC CACGGGACGC CCTCCTATTT GCCGTGGACA
 2051 ACCTGAACCG CGTGGAAGGG CTGTACCCT ACAACCTGTC TTTGGAAGTA
 2101 GTGATGGCCA TCGAGGCAAG CCTGGGCGAT CTGCCACTTT TGCCCTTCTC
 2151 TCTCCCTAGT TCGCCATGGA GCAGTGACC TTTCTCCTTC CTGCAAAGTG
 2201 TGTCGCATAC CGTGGTGGTG CAAGGGGTGT CGGCGCTGCT CGCCTTCCCC
 2251 CAGAGCCAGG GCGAAATGAT GGAGCTCGAC TTGGTCAGCT TAGTCTGCA
 2301 CATTCCAGTG ATCAGCATCG TGCGCCACGA GTTCCCGCG GAGAGTCAGG
 2351 TGAGAGGAGC CTGGTGCGTG GAGTGGAGAT GGGCGCTGCT GGGGGCCGGG
 2401 GCCATTGCAT GAGGGGAGAG AAAACGCGTT GGTAAAGTCT GAGGGGAGTT
 2451 GTTACTTTAT AACTTTGATA TTGCTTAACG ATTGGGCCAT GTTCGTAGGT
 2501 GGTAGGTAGA AGGAGCTTAG TAGAAGTAGA ATAAATATT TAAAGCGCGG
 2551 ATGGAATAAA AACGCGCAGT GAGGTCGCGG CTGGAAGGAA AGAAGTGGGG
 2601 AGAATATGAG AGAAAATCAT ATTTTGACCG GCTGGGAGAA ATCTAGTAGA
 2651 TGCCCGACGG GAAGTAGAAG TCGAGGTTCA GGACCGTGGA GAGCGGTGAA
 2701 GGTTCTGAAG AACTACAAGA GCAGGGTATG GGGGTGGGGT ATCCCTGACT
 2751 CCTGGCTAGG TGTCACACTC CCAAGAGCAA CTCTGACAGC ATGTGTCGGA
 2801 AAAGCAGCAT CTGCTCTCTC TGAAGTGTG AGAAGGTGTG CCTGAGCCTT
 2851 AGGCAAAAGT GTAAGGAAGA AAGCACATCG CTCTGAATTC CTCTGGGTAA
 2901 ATAGAAAATC TGCACCTAGT ACAGAAGCCA TAGGTAGAGA AGAGTGGTCA
 2951 ATTAGTCTCG GATATTGGAA AGCATTAGAA ATATATAAAA GTGTAAAGAT
 3001 GGACGGGGAG ATTTATTTGG GGATTGTTTC TTTGTCCCTA CTTTCTTCT
 3051 ATGTAATGTG GACTCAGAGG CTGGTATTCA GTTGCTGTGT TCAGCCATT
 3101 TCTCTCCCA TCATCTAAGA ATTAATAAAA AAAAGGAATT AAATGATTTA
 3151 GTTCTTATT GATTAAAAA GCTAAACATA TTTCAATGA AAGAGCTATT

FIGURE 3, page 1 of 27

```

3201 TGTGAACTTA ACGTTGACAA GTAATAATGA GGAGATGAAT CTTTAAGGAC
3251 AAGACAGAGT CCTTATTTAG TAATGAGTTT TCTGCCTTTT ATATGTTACT
3301 TTTATCATAA TCTCAAGCTG TGTTAAGCCT TGCACAAAGT ATCTATGAAG
3351 CAAATAGGTA ATTGGCATGG GCCCATTTTA ACGACTGAGA AACTGAGAAA
3401 AGTCTGGGGC CTTGATCAAC ATTGGTCAGT GTACTAGTAA AAAAGTCCAC
3451 ATTCTGGAGG ATTTCTGACT CTACATCATT TTCACTCAAC TCTTGCAGCT
3501 GGAAAAATAT GTATTTCCAA TCTTCTTCCA CTTCTGATAT ATGTGCCGAG
3551 ATAAAACTA AAATGAGTAA GGGCAATGTA CAATGAAAAG TTTGATAGAA
3601 TCTATGCATA AATTGTCAAG GGAGTACTAA AGATTCTTTT TTCTAGAAGA
3651 AAATAAATCT TACATTTTTA ATCTTAGGAA GGTTGAGTAC AAGCCATATT
3701 CAGCAGTTGC CCAGAAGATT CCTAGCCGAA CTACAGAGAT TTGATCTGTA
3751 GAGTGCAGGC TAATTAACCT TTATATAAAA TATTTCTGTC ACCTGAATCT
3801 GAATGGTGTA GGCAGCAGAT GGAAGGCAT GGAAAAACA GATACACAAT
3851 GCTGCCAATG ACCAAAGTGT TATAAACATG AAATTGCATC CATAGGGTGC
3901 ATCATTATTA ATATACATGC AGAATCAGAT CTAACAAAAT GCAGGAGTCA
3951 GCATCACTTG CTTCTTATCA TTGCTTCTTT ATTACCTAAT ACTTCGTAAG
4001 TGGCCAATAG TGGTCACAGT CTCCAGACTC CTTTTCATTT GTAGATTGTT
4051 TGGCAAGAGT CTTAAGTAGC AAGAATTTT CACCAAAATT CGTGTTCCCTT
4101 AGTTAGAAGG GAAGTTGTGT TCTAGAACGG ATGTGTGCAG CATACAGCAC
4151 ATTACAACAA GGGGACCAGA AAAACCATGA AGCAGAATCC AGAATCTGTA
4201 AACTTCAAAG CTCTAAGACC GGGGTGGGGG GTGGATAAAG TCCTCCAGGG
4251 ACAAGCTGAC AATAAAATAA ATACGCTGTC AAGCACATTT GTTCTGTAT
4301 TTCAACTCAG AAACATATTT TAAATCACTG TTGTCAGTGT TACCTTCATG
4351 GCACACATCT TGAAAGGGAG AGATTATTAT ATTAATCAG ATCTAGTTTG
4401 TTCAACTGAC TACATTTTCT TTCATCTCCC TTTTACTTT AAATCAAAC
4451 ATATCTAATT TGCTTACTT TTGGTATCTC ATTTAAATGT CATCTATAA
4501 TATTCTGTAT TAAGAATGTT CTGATCAATG CCAAGTCACT GTAATATATA
4551 ATTTTAAGAT GACCGTAATC TGCTTTCAGT GAAAACAATA ACTGATCTTT
4601 CCCTTGCTTC TCTGGAAAAG TGGACCTTCC TCTAATGCAG TGATGTGATT
4651 TTTAAAAACT TTCTATATAT AAAAGGATGT CAAACTCATT TTACACATTA
4701 AATAAAATG ACTTAATCTA GCCACCATCC TCGGAGTCTA CTGCCAGTG
4751 ACCTAAATTG TTGGTTGTGT GCCACTCCCT GAATAAAGGA TTCGAGAAGA
4801 AAGTGGACTT TTTACACAA CCAAGTAAAA TAAAATGTG TCTCTACTT
4851 AAATCAAAAT TGCTTTCATA GCAAGAGCAG CAACAGCTGT TTTCTCTCAC
4901 TTTATTTTGG GCTGCTGATT ACATTATCT GAAGGTTTTA ATTAATGAGC
4951 AATAGTTTTG GTAACATCCT GCCACAATC TTAATGGAA AGAGCTGCCA
5001 CTGAGATGGA CAAACCCCTG AGAAAAGCAT AATAGTTTTA TTTCAATACA
5051 CTATGTATTC AAAATAAGAT AATCACATAA GATCATCACC TTCTAGGGGA
5101 TCAGCTTCTT TCAAGTAGGG AAATTCATTA AAAGTAAGTT AGTTAACTAC
5151 ATACTTTTGG AAAACATATG TATATTATAA CTGCATAATA AAAGCTTAAT
5201 AAAACATTAA ACATAGGATG GGGTCAAAGC AGTTTTCAT CAAAAGAATT
5251 CTGACTTCAC TACAACATC AAACCCCACT TGAGGCTAAC CCATTTTATT
5301 AAAATGATTA CTTCTTTGTT CTAAATTCTA TTCTTATAAC CTTCAAATAA
5351 TGATGCTGAA TATGAACCTA ATTCCATTTA CACCTAAAT AAATTCCTGC
5401 AGTTATACTT TCTTCTCTC TCCTTCCTAC ATGACTTTTT CTTCTACAGG
5451 TTTTGTGTAG TTTCTTCAA GTTAACTCCC TAAAGTTTAC CTGCTGAAGT
5501 AGTGACAAAGT ACACATTTTT TTAaaaaaat ATACACCTCA CCTTAACTTC
5551 ATATTGGTTC TATTAGGCAG AGTTAATGAT GTAATATAAT TGGCTTAGAT
5601 CCAATCCAT GCAATTCAAA AGTGACTGCA CAGCCAGGCA TGGTGGCACC
5651 TACATGTAGT CTCAGCTGCT TGGGAGGCTG AGGCAGGAGG ACGGCTTGAG
5701 CCAAGAGTT CCAGACTACC CTGGGTGACA TAGTAAGACC CTGACTCTTA
5751 AAAAAATTTT TTAATTAaaa AAAAAAGTGA TCCCCTACT ATTTTCAACA
5801 CTCTCGTTGA ATACACCAAC CACAACTTG CCTGCTTCAT GAGCGATATG
5851 TACTAACAAA TTAATATATG CTTCTTTCAT GGAAATACAA GTGTTTAAA
5901 TTGTGCATTT TCTCTGACAG TACAGGACTA AGCACTGAAG CCTATTATT
5951 AGAATTTGGC TAACAAAGCA CTATTTTTCG ATGGCACAGG GGTACCTCAT
6001 GAGGGGACCA GTAAGGGATA TTTATTTTAA AAACATCTGC TCTCAACTGT
6051 GTTGGTTTTT TTTGACTTGC TCTATGCAA TCACAGCTCT TTCTCCTCTG
6101 GGGGAAATGT ATTCTGCAAT TCATGATGAA TAGCTGATG TCGCATCTAA
6151 TTGTTGCTGA CTTAAAGATA AACAATTTCA AATTAAATGT CAAGTGATGC
6201 AAAACTTTTT AAAGCAGTGA TCTTTTACAG GTTCCTCTTG AAGAACAGAG
6251 ACCTGGCATT AACTTGGAAG TATTTTAA TTTAGTTATT TACTTACAAT
6301 ATGTATTCGC TTTTCTAGAT AAGTAGAGCA AAGGAGACTA GCAGGCACCA
6351 TTTATTGAGC AATTAGTTG TCTCTCCCGC TTTACTTTGT GCTTGCCAGA

```

FIGURE 3, page 2 of 27

6401	AGAGTATCAC	TTAATCCATG	AAGAATATAT	TTGCTCTTGA	TTTTGTCTGA
6451	CCATTATATC	TTAGAGTTAA	TTTATGATCG	AATCAGCTGA	GGTATCTGAA
6501	GACTGATGCC	AATTTCTAAT	TCCTCGTGTT	TTATCTTCTG	GTGCTGCAGA
6551	AGGCACCATG	GATTTTGTAC	CATTAGATTT	TATTTTATAA	ATAACCCCCC
6601	AGTCAAATTC	CAACCACAAT	AGTTAAAAGA	GCACAATGTA	ATGAAACGCA
6651	TATGAAATAG	TGGCCAAAAT	GTTCCCAATC	TGCCCTCTTC	GTTGCAGATG
6701	TTCCCAATCT	ATCTCTCACT	AGCCATGTAA	TTTTGGGCAA	GTGACTATCT
6751	CTGAGCATCT	ACACTTATCA	GAAGTGTAAG	TCAAAAGAAG	GTACTTTACC
6801	AATCTGACAA	AGATGTGTG	AATCAAATGA	GAAAGTAAAA	GTACTTTGGA
6851	AAAGTTATAG	TGTTACCCAA	ATAACTAGTG	GAGAGGTGTT	GGTCATTATC
6901	TGGTAAGAAT	CACTTAAGTG	TTAAAGTTCA	ACTAATTTTC	TTTTCGAAAT
6951	TACTAATCAA	ATGAGATAT	GATTCACATG	TAAAATGTTT	GCATTCGCCCT
7001	TACACCTTCC	TCTTTACCCT	CTCCCTCATT	TTTTTCTCTT	AAAAAAAAGT
7051	GGGCTTAGAA	ATAAACAAAC	AAACAAAAAA	CTAGGTTTCT	AAGTAGGTTG
7101	CACACTTGCC	TGGAAAAAGA	AGACATGCCA	CAGTACTGTT	TCTGTATTAC
7151	CAGCAACTTT	TAACCTATGG	CACATCTAAC	ACAGCTTCTA	GAGCCTTAAG
7201	TCCTGCCATA	GAAATATCAT	TAAGATGCCC	AAGATATTTG	AGAAATGTTG
7251	GTCTTTCACA	TTGCTCATAA	GTTTTTTCTA	TAGGCAAACT	ATCATTTCAGG
7301	AAATTATGAC	CAAACAGAGT	CTACCCCACT	CTCACTCCTA	TTCCGCCAAC
7351	TACACCACAA	AGCAAACATC	CAAATTTTTT	CATAGCAAAAC	TTTCTTGATA
7401	AGGAAAGCAG	TGTGTTGATT	CATACTGACC	TAAGCTCCTT	ATCTCATCAT
7451	GGATATATAA	TTTACAAACC	AGCTACTTTG	AGTCCCATTG	CCCTAGATAA
7501	CTGTATACTC	TCTTAGGAAA	GTATTGCTCA	TTTTAGTGGC	AACAGTAAAT
7551	ATAGAGATGA	GAAATCTCAT	TGTCTTTTTT	TCTGCTAGCT	CTGGCTATTG
7601	CCACATATAC	ACAAGAATAG	AGGACCTATG	TAGCACCAGA	AATATGATGC
7651	CAAATCCATA	AAACTAGGCA	AGAAGAAAAG	ACATCTCTTA	GCATCTGCCA
7701	TATTACTTTT	TGAGTGAATG	TTTGAATGAC	AAACTCATAG	AAATTTTTAA
7751	CCTCTCAGTT	GTCTTTTGCT	GATATTTTCT	CTATGAGATC	ACAGGAGCAA
7801	AGAGCAATGG	GGAAGAGGTG	AGCTAAGAGT	AAGCCAACTC	TCTCCATTTT
7851	CTTCTCCTTT	CCTCAACCTG	ACACCTCAGC	CAAATTCCTA	GATATTTTAT
7901	TTAATTGTGT	AAGATTCATG	CACTTTCAGG	GAGACTGTCA	ACTGTATTTG
7951	ATCTCCTGCC	TTGAAAAAGT	GGGTTGATCC	CTGACCTGGC	TGGCATTATG
8001	TCATGAGGAG	TAAACTTTGA	CATTGAAAGG	CTGGTGTCTG	ATTGGCCAGG
8051	GCTGGTTTGT	TAATGGAGAT	GAGCCTGCAA	GGGTTTCATG	GGTAGAGAAT
8101	TAAGCAGGAT	GACTCCTTTT	TCAGTCAGGA	TGAATTACAC	CTCATCGTTA
8151	TTCAATTGAA	GCATAGATTT	AGGAAGAATT	TTAATCATAT	TATATTTTTG
8201	CAGCATTTAT	GTTTTCAAGG	CTTTTCCCCC	AAATATGTTT	AAATAATCTC
8251	CTAACAGCCC	CTGTAAAGCA	AATGACTATG	TATAGACAAA	ATATGGAGAA
8301	ATAGAAAAGAA	CTATGCATTG	TAGAAAGTGAG	AAGCCATCCA	GAAGAAAAAC
8351	AGACTAATAG	CACCTTCTTA	ACTTCGTCTA	TTTGTCTAGC	TTCAACCAAA
8401	GAACAGATTG	GATCCATCAC	AGCTTTTGTA	ATATCCCATC	AGAAAAGGCA
8451	TAAAAGAGAG	AAGTATCTAC	TTCCAGGTTG	GACTATTAAA	AGCATATTAT
8501	ATAATAATGT	TCAAAATGAT	GGAAATGTAA	TTAATTGAAT	CATTTTTTAA
8551	CGGAGTACAA	TGTACTCATT	CAAAATGAGG	ATATATTTAT	TGACACAGAA
8601	AGATAGAAAA	ATTTTACAAA	AATGTATGTA	AAGTGTGACC	TTGTTCTGTG
8651	AAAATATGAA	GTCTAGAACT	TTGTTTATAA	AAATCTTAAC	AGGGATTATA
8701	GAGGGTGAAA	ATCTACATTT	TTCTCTATTC	ATGCATCTAT	TTTTCTAAAA
8751	ATAATCATTT	TTACCTACAG	AATTAAAAAA	TAATAAGAGA	AAAAAGATAA
8801	GTTGTCTTTA	TACTTTCTCA	TCTAGACTTT	CAATACTTCA	GTATGAAAAT
8851	ACATATTTTA	GAAAGAATAA	ATTGTAGTGT	ATGTTTAAAT	GTATCAAAGC
8901	CCTCTCTCAT	CTCTTTTATA	ATTTTTTCCT	CACAGTTACT	TTGTGAGGAG
8951	AGCCAGAAAA	ATATTATTAC	CATTATAATG	TAAATGAGAT	AATTATAACA
9001	TATGATGGAG	GAAATATAAG	GTGGGAGTTA	AGAGATTACA	TTTTTGAGAA
9051	GTTTCAGAGCT	ATGTTAGAAT	CTTGGCTGTA	CAACTTTTTA	GACATCCTAC
9101	CTTGGAAAAC	ATACCTGATG	TCTCCATGTT	TCAATTTATT	AAGTCATAAA
9151	ATGTGATTAT	CTTATAAGTT	TATGGTGAGG	ACTGCACCAT	ATGTACCTAG
9201	AGCAGTGTGT	TCTTATGATG	TTGATAATTG	TTCTTTTAGC	ATAAAGTGTT
9251	GACGTGGCAA	GATTCATAAA	AACTAATCAG	AAAAAGAACT	CAAAATTTCT
9301	ACCTTATTAT	GCATTGCGAA	TGTGTCTATA	ATATTACGGC	TTAGGACTTC
9351	CAGCTCCAGC	AAAATAACCC	TGAGAAAAAT	GAAGAAATCT	GCTGTTTTGA
9401	AGTCCCACTT	AGAGTTCTGG	TTCACTGAAG	TGTACCCGCA	ATTTAAGTGT
9451	GTGCAAAGTA	GGTCAGCAAA	GAAGTGAAGT	TTGAAGTCCA	GTTTTACCTA
9501	GTTGCTCCTT	TATATGGGTT	CAGGGTGGTT	GGAGTTTTCG	AGCAGTTACA
9551	TCAAGGTTAA	GAAGAAGCAT	GTTTTGGTCT	ATTAGGTGGT	CTTAGTGAGG

FIGURE 3, page 3 of 27

9601 AACTCATAAG TCTTTCCTAA CTATTGCTAT AACTTCTCAT AGGAGGCTCT
 9651 GAGGAACTAA ACTCAGGGAA CAATAGAACA GAAATGACAG TTTCATTTTA
 9701 TTAATAAATG CATTAAATGCC CAGTGCCCTG CTGCATAGGT CTTTAGAAAA
 9751 AATTGAGTTG GGACATACGA CTTGGGCTTC AGGTTTGTGT GGCATTTCTT
 9801 AATTCTAAAT CTTGATCTTC CATCTAAGCA AACAAAAGAA AGAAGTGGCA
 9851 GAAGAGATGG AGGACAACAG ATATGAGCTT ATGAAACAGG AGTGAGCTTA
 9901 TTTTGGTGTG GTAGGGCTGA GTACCTGGAA GAGTTCCAAA TCTGAATCCT
 9951 CAAAACCTGT GAATATGTTA TTTTATATGG CAAAAGGAC TTTGAAGATG
 10001 TGGTTATGTT AAGCATCTTG AGATAGAAAG GGTATCTTGG ATTATCCAGG
 10051 TGTACTCAAT GTCATCACAA GGATTCTTAT AAGAAGGAGG CAGAAGAAGT
 10101 CAAGGTCAGA GGACAAGGCG ATGTGACAAT GTGATAAAGG AAGCGGAGAC
 10151 TGGAGTGACA CACATTGAAG ATGGAGAAGA GGCCATAAGT CAACAAATAC
 10201 AGGCAGTCAC TAGAACTCA AAGGCAAGAA AATGGGTTTT CCCCTCAGAG
 10251 GCTCCAGAAA GAATGCCACC CTTGACTTTA GCCCAGCAAA ACATATTTCA
 10301 GACTATGACC TCTAGTCACA AGTAATAAGA GAATAAGTGT GTGTTGTTTT
 10351 AAATCACTAA GTTCTTGGTA GTTTGTACA GCAACAACAA GAAATGAATA
 10401 CAATTGCCCA CACAGACTTA TGCAGGGAGG AGGTGACAAA AAGATAGAAA
 10451 GGGATCTGGC CCACCTTATC CTTGGAAGGC AGGTTCCATT TCAGCTTATA
 10501 CTGATTGCAT CTGGAGAAAA ATGTCGAGAA AGGTTAAGCT TGGTGATTCC
 10551 ACTCACTGCT AAGTACAACC CAGTTGGTAT TTGGGGATTG TGTTAGAAA
 10601 GGGCAGAGTC TACCTGGACG TGGATTCAAG GTTCAGCAGT CTCCCCTTTT
 10651 TCATACGGGC TTCATCTGTA TCACAGTAAC AATATGGCTT ACATTAACCC
 10701 AAAGATTAAAG GGAAAGTAGC ACTGCTATAG GCCAGGGCTT TCAGAAATGG
 10751 GACACCTCAT GAAGCAAAAT CTCCATATTT TACTGGATGC GGACTATATC
 10801 AAAATTGATG CACACAACCT CATGGGACTT TCTAGATGCA TATTGCTTTT
 10851 CTGATTATAA AAGCAGTGCC CATGCACTAC TACCTGTGCA TTTACACAGA
 10901 TTAACCCTGA GTCGCATTTA ATGCTTTTTA TTCTTTCAAG GATAATGGTT
 10951 GAAATTTTAG TAACAGTGGA GTACAATTAA GAAAAACCCT GTTATCACTC
 11001 TAACTGGGCA CTGGCATCAA AGAACAGGAG AAATAAAGAA AAAAGTAATT
 11051 TTTAAAAAAT TTTCTGAAAA TGCAGATTTG ACATGGCTTT TGAAACTTGA
 11101 GCATTATGCT ATTTCTATTT AAAAGGTGAG ATTTTCCTTT GTGTTTGCAG
 11151 TCTATATTTT CATCACACTG CAAGTGGCTG AGTCTCTACG GTTCCAAACA
 11201 ATAGCTCAAC TTGTACCTTT CAAAAACATT CTTAGGAATA ACTTAGAAAT
 11251 GGGTTGTGAC TCCTCTCCTC ACCGCCAGGG GTGGTCATTA GCTGAACTTA
 11301 CTGAACATTT GGGGCAGTAG CAAGCACTTT GATGGCAGTA CAACCTGCAT
 11351 GCAATCTATG GGTGTTTTTG GACAGAAGGC CTCAACTAGA AGCCAAACAG
 11401 AAGTTGTGTT AATACTCCCC AGATTAAAAA GAAAAGTTTT TGTTTTCGTA
 11451 AAGTTCAACA TTCAGCATGT CTTTGTCTAA CAGAATCACA ATCTGGCTTA
 11501 GTTGTGGAGT GCTATTTTTT CAGTCCCAAC CAGACATTCT TAAACAGAGA
 11551 TTCCTTTAAA CAAATAATTT GCTTCTACAT ATTGTAAATG TAATAATGGG
 11601 AGCAAATATA TACACAGATC CACACACAGA GAGATGTTAT TGTGTTGCTG
 11651 ATACAGGAGG AGTTAATTTG AGTCTTTTCA CACATTGTGT TATACACATA
 11701 AAGAAATGCT TCAATGTGAC CTGAACATGA ATGATAAATC TAGATCCGAA
 11751 TTTATCTAGT GTGCCCTTAC CTGGCCACAG ACACAGAGAG CCATCTAGTG
 11801 GTCTCCAAAA TACAGCTTTA GGCTGAAGCA TCCTAGGAAT CCAGTCTCAC
 11851 AAGACAAGAA AGGATTCCAA GCAGCTATTA CTTCACTTCT GGTCTTTTGA
 11901 CTGTGGAAAA TGTAGATTAA TTCACCAAAA AGATCTTCTT CTGCCTTCTA
 11951 CTAAGAAGTT TCATCAACTT CTGCTGTACT GCCAGCCTAT CTATAATTGC
 12001 AGTTAACAAC TATAAAGTAA GATATCTCAA AATGTGTCCA GTGGGGTTGG
 12051 GAGAAAAATG AGATAAAGTC TTAATAACTT TAGAAATGTA GAGTCATTAA
 12101 TTCTTAGTAG CTGTATTTGC TGTCACTTTC ATTCATAAGG AAAGATAAAG
 12151 AGATGCAGCC ATTTTATTGT GCTAAGCACA TCATTTATTG CTTATTCTCT
 12201 GATTATAAAA AAATCTATGC TCTTTGTGGA TAATTCTAAG ATTTAATAAA
 12251 ATGCCAAAGA ATCCAAATCA CCTATAAATA CACAACCAAA AGATCTACTC
 12301 ATATATGTGT GTTATAAAAT TGGGGCCATC CTACTTTATG GATTTCCTTA
 12351 ACATATTCAT ACATTATAAA CCGCTTCTCA TATCGCAAAT ATTTTCTCTC
 12401 ATTATTTGTA ATGGCTAGAT AGTTTGCTT TGAATGGGTT ATTTCTTGGT
 12451 TTGTTTACCT AGCCCTTGCT GTTTAAATC AGATTGCTT CCATTAAAAA
 12501 AGAAAAAACA CTGTTACCGA AGGGATTTTA CTACACTCAT CTTAGCATTT
 12551 TTGTAGTTAC TTGTGCTGCA GAACACCCTC TACTTGAGTT TTGTGACACC
 12601 GTCATTGTTT TCCTTGTTTT TTTCCCCTGT CCTTTCATTG TCTTTTACTG
 12651 ATTATGCTTC TTGTCCTTGA TCCCTTATAA TAATCACCAT TGTACTGTGC
 12701 TGTCTAATAT TATAGCCTCT ACCCAAATAT GGATACTTGA ATTTAAATTA
 12751 AGATTACATA AAATTTGAGT CAGTTTCTCA AGCACATGAA TCACATTTCA

FIGURE 3, page 4 of 27

12801 AGTGTTTAAA ACATCTTATG TACCCCATAA ATATATATGC CTACTATGTA
12851 CCCACAAAAA TAAAAAATT TTTAAAAATT TAAATTTAAA TTTAAAAATA
12901 ATAAATTTTA AAAATATAAA TTTTAAAAAT TTTTAAATAA AATATTTTTT
12951 TAATTTTAAA TAAAAATTTT TAAATGGGAA CATATGGGTA GTGACTTCTA
13001 TATTGGACAA CATAGATATA GAACATTTTC ATTTCCATAG GAAGTTCTAT
13051 TGGATTGTTA TGTCAAAGGA TGTCTGTGTG CAAAGGAAGA GAAACTCCCA
13101 CCATGAAGCT CAAAAGCAGG GAATTTGTGC TGAAATCTTA CAGGAAAATG
13151 ACATGAAATA GAAATGCATG AAGTATAGCT GAGCTGTACC ACCAGAAGGT
13201 GTTTGGGTAG CACCTCTCTT TTTCCCTTGC TCTGGGGCCC AATGGCTCTT
13251 CTCTCAGTTT CTCACTTCAC ATCTGCTACA AACTCCTCTC GGGATACCAG
13301 CTGATTCTTC TGCCTTGCCA TAGCTTCTCC ATGGATGTGG TTCTTATGAT
13351 AGGACTTAGC CTGACTCTAT ATGACTTTAC AACTCTAATC AATTTATCTG
13401 ACTACATTTT TTATATCTCT TAGTTCAAAT TATCCAGATA TCTGATTGGT
13451 TTAACCCACA TTGGTTTCCC CTCACCCCAA ATTATATGTC CTCCCTTAAT
13501 CCAATCAGAA AGGCCCAGAT TCTTAGGTTG CATGCTCAAT ATGAGACTGA
13551 CTGAAAATAC AGCAGTAAAT AAAATAGCAA AGCCAGGCT CTCAAAGAAC
13601 TCCCATTCTA GCAGGAAGA TAGAAAATAA GACATGCAAA CAAATAAATA
13651 CGTAATATAT TATTATTAGA CAGAGACAAG TGCAATGGAG AAAAATACAG
13701 CAGAAATATA GATTAGAGAC TGACTACAGA TGGTAAAGTA AGGTCTGTCT
13751 GACCATGGCA TTTGAGCAGA GAATGAAGTA AGGAGTGACC CATAAAATCT
13801 TCCAAGGAAG GAGCATTGCA AGCCAGGCCA ACATAAAATA AAAAGACCCT
13851 GAGATAGGAA TGAGCACAGT AATAAATAAT TGTGAATAA GGGGACTATT
13901 CTTAGTACTA TCCATAACAC AATTTTtagT GGGCACTGTT TCAAAGGAGG
13951 TATCAATAGT GAACCAATAA CCAGATCTAG TACACCCTTT CATACAGGCC
14001 TTTTCCATAG TGTCAACTAC TGAATTTATC TCTTTGTGTG TGGCAAGGCC
14051 AGGAATTTCT AACTTGAAAT TGTGGTTATA TCTCCAATTC TCACCTTAAG
14101 TTAAAAATAC TTAAAGATGT CTTGAAAAAG TGTTTTCTC TTACCTATAA
14151 CAAGACTTTT CATAACATCT TTGACTTCTC CCTTTCTTG TTACCAGGTT
14201 CTGTTGCTTT CCTTCATATA TTTCTCATAG CCCCATTCTT CCTTCTTATT
14251 GTCACATTAC CTTCTTGCAAT CAATTTCTTT GAATAGCCTT TTAATATCTA
14301 GCTTCTTTCC ACCAGACCAT TCTGCACACT GCTGCTAGAT AAATTTCTT
14351 AAAGCAATCT TTTCTTTTTC ATTCAATCAA GAAATACTTA TCAAATATAA
14401 TGCCCTGAGC TTCATGCCAT TCCCTTGCTC AAAAATCTT TTAATATAAT
14451 AATATTCCCT TTCTTTTCC ATGACCCAAC ACTTCTGTGG GGTGAAATAC
14501 ACACCTTAAT AACAATGACT CACTACAGCA TTAATTCACA AAATGGAGT
14551 GGGGTGTGCC AACTCAAGA AACTGTATTA AATTATCTAG ATTTTGAGAG
14601 TATAATTCAA TAAAGCATT CACTCTCTAC TGACATGCCG AGATTGAGC
14651 ATGTTCCCCA TAAAGCCAGA GAAATATAGG TTAATAATCA TCAGCAAGTT
14701 ATAGAATCTG GCCCTCAAGG CCATCCACAA ACATGTACTG TATTAGTTCA
14751 TTTTCACATG CTGATGAAGA CATACCCAAG ACTGGGAAGA AAAGGAGATT
14801 TAATTGGACT TACAGTTCTA CATGGCTGGG GAGGCCTCAG AATCACAGCC
14851 GGAGGTGAAA GGCCTTCTT ACATGGTGGC AACTAGAGAA AATGAGGAAG
14901 AAGCAAAAGC GGAAAACCCC TGATAAACTC ATCAGATCTT GTAAGACTCA
14951 TCTACTATCA CTAGAATAGA ATGGGAAAAC TGGCCCCCAT GATTCAATTA
15001 CTTCCCTCTG GGTCCCTCCC ACAACACATG GGAATCCAGG CAGATACAAT
15051 TCAAGTTGAG ATTTGGGTGG AGACACAGCC AAACCATATC ATGTACTCTT
15101 TCCAATTCAT GGCATTCTGT TGAGATATAG GTACACAGAA AGCACAGAAT
15151 TTCTTTTGTT TTACTTCTAT TTTAAGTTCA GAGGTACACA CGCAGGTTTG
15201 TTACATAGGT AAACTTGTGT CACAGGAGTC TGTGTACAG ATTATTTTCT
15251 CACCCATGTA TTGAGCCTAG TTTTCATTAT TTTTCTGAT CCTCTCCCTC
15301 CTCCCCTCTG TCACCCTCTA GTAGGCTCCA GAGTCTATTG TTCCCTCTA
15351 TGTTTCCATG AGAAAGCACA AAATTTCTAG AAACAGAAAT GTGTGTATGA
15401 TTTTAAATC AATACATATA AATCATTATA TTAAGGCA TTTTCTATT
15451 ATATATCTAT ATGGAAAGAC GGATATATAC CCAAGTTGTC ACAATTTGCA
15501 GATGAATTAT GCTCTAATTC AAAATTGATT TTTCCATTGA AACAATGTGA
15551 TCTGTTCTTG TTAAGACCTC AGGCCAGGCC TCAAAAGCCT ATTTGACCCA
15601 TTGTATAGCA GAGTTCTGGT ATTAATAATT CTATAGACAC TAAACATCAT
15651 CTGTAACAGA CTCTTCTGT TTAGACCAA GGGGATATGG AGTCGGGAGG
15701 AGAACCAGAG ACCTGATTTT AAGTTTGGTT TTAGAATCAT CTGTAGAGCT
15751 TTGGGAAACT TCTCTGAGCC TCAGTTTATA AATAGTCATT CATTAACTG
15801 GTTTTTATTG AGAGCCTACT GTGCCATTTA AAAAATTA TACAGACTTC
15851 AGTGAATTAA TACACATAAA AGCACTTTAT AAATTCAAAT TTTAAAAATA
15901 GATGAGAGGC ATTTGTTATTG AAACATCTTC AGGAAAACAT ACTCCTAGCT
15951 TCAATTTCTG AAAGTTAGGA CCTATCTTCC TTGGTACTAA TTTGGCAACA

FIGURE 3, page 5 of 27

16001	GGAACAACCC	ACCCTTGTTT	CATCCTCCTG	CAATGGACCA	ACACAGTCAA
16051	ACTGTAACCT	CTAAATGGTC	AGCAGCAGCT	GGAAGGGGAG	GAAAAAAGAG
16101	CAGGGTTTCA	TAATTTCCCA	ACGGGGACTT	AAAAAGTGTG	TTTATCTTGG
16151	ATGCTCCCAT	GGTCAGGGAG	AAGAACCCAG	GGTGCTCGGG	TGTTACCTTT
16201	AGGCCTGAGG	AGGAAGAAGG	GAAGTTGGGG	AGCCATCAGG	ATAGGAGGAC
16251	TACAGCCAGA	ACACAGATGA	GAATAAGAGA	CAC'TTGGGAA	GTCAAGTATT
16301	AAAGCTAGGA	TTGCTAGTTT	ATATTCATAA	AAATATATTA	GTTAAGATTT
16351	AAGATTGCAT	CAGTTTCTAA	ATAGTACTGG	GTAGTGGGTT	GAAATACTGG
16401	AAATGATCAT	ATCCTATTCA	TAACCTATGA	AGCTTACTTC	ATTCCAACCTC
16451	TGTCTTTAAC	ACTTGCAGGG	CAGCAGCCAC	TTAAAGTCCT	TTGCATCTCC
16501	AGCTTTTCAG	ACTACTTCAG	GATTTAGCCC	TGAGCTCAAG	CCAGGGGAAC
16551	CATTAGGTTT	TCCTTGACAG	ATGAGAGGGG	GAAGTAACCT	TAGGAGAGAT
16601	CAGTAATAAA	TCAGTAAGCT	TAACCATGGC	CATACCATCT	CTGCCTACAG
16651	TATTTCAATG	GCTCCTAACT	GACTTAAGAG	GCCATTGAAA	CACTGAAATT
16701	TAAATGGCCT	CCTAACCCAT	CCTTTACCAC	CTTTTTTTTT	TTTTTTTTTAA
16751	GATGGAGTCT	CACACTGTTG	CCTGGGCTGG	AGTGCACTGG	TGCAATCTCA
16801	GCTCACTGCA	ACTTCTGCCT	CCCAGGTTCA	AACGATTCTC	CTCTCTCAGC
16851	CTCCTGAGTA	GCTGGAATTA	CAGGCGCATG	CCACCACACC	CGGCGAATTT
16901	TGTCGTATTT	TTATTAGAGA	AGGGGTTTGA	CTATGTTGGC	CAGGCTGGTC
16951	TCAAACCTCT	GACCTTGTGA	TCCGCCCGCT	TCGGCCTCCC	AAAGTGCTGG
17001	GATTTTCAGG	ATGAGCCACT	GCGCCCGGCC	CACCACTTCT	ATTCTCTTGT
17051	CCCAGCTTCT	GTCAGAAAAA	GAATCGGTGT	ACTAACCTGC	TTAAACCCCT
17101	AAATGGCAGC	AGTATGTCCC	AAACTTCAGG	CATTTCAGTTA	CTGCCCTCAT
17151	AATTTTTGCC	ATATCTCTGC	ATCATCTACT	GCTATTAATG	TTTTAAATTA
17201	ACTTGTTGTT	TTACCTAAAT	AAATTCACCT	TTAAAAATCT	TTTCATGACA
17251	ACATTAATGA	AATACCAGTA	CCATTGCGCA	TAAATAGAAA	TTAACTATAA
17301	AAATAAATAC	ACAATAAAAA	CTAAACAGTT	CTAGCTAGGT	ACCCTTGCCT
17351	GCCTGAGGTC	TGAATCTGAG	TACTTTTTTTA	AAAGAGGAAA	TTTCTAGGTG
17401	CTATAAAAGT	GTTAAAGACA	CGCTGACACC	AAACTGAGGC	TTTCTGCTTA
17451	AGTAAACAGA	TGGATTAAAT	GCTAATTGAA	AAGGAATTAA	GTTTCTCACT
17501	ATGTGATTCA	GTGTTATATT	AATGTAAAGT	TTCTGAACAA	CCTAAAAATCA
17551	TGTCATGAAT	CACCTACACT	CTGCCAAACA	GTAACCTATA	AGGTGAATTC
17601	TAAGCAGCTT	AGCGTAGCAT	TCAAGACCCT	TCATTATCTG	ATCCTCACAT
17651	CACCTCTCTC	CTCATTTATT	CTTCATACTA	ACACTTGCCC	TTTGTACTTT
17701	GTGCTCCAGT	AATGCCTAAA	TGTGGAATAC	TATTCCAAGC	ATATGCACAT
17751	GTTGTTCTCA	CTGCTTGGCA	TACCATTTTC	CCTTGTGTCT	GCCTGAAATT
17801	CAATCTTCAT	CCTTTGCTCT	TCTGTGCATG	GTACACTGGC	CACTCTCTCC
17851	CTACCATGAT	TGACAACTTT	CACCTCTATG	TGACTTTTCT	ACGGTCATCT
17901	TTCTAGATCT	GTCATACAGT	TATGTAATTA	TTTGTTAACA	TGTGTCTCTC
17951	TCCTCCTCTC	TCACTAGACC	AAACCCTGTG	CTCCTCACAC	AATGTCTGGC
18001	TCATAATAGA	TGCTCAATGA	CTATTGGTTA	AACTGAATTA	ATGGTCCACT
18051	TTCAATTCAT	CTAGTGTAA	TGCTAAATCA	CACCTGTGGA	AAACCCACCA
18101	TATGTCAAGG	TATGGTGATG	GGAACCTAAA	AGAGTGCAAG	GCCCTGTGAA
18151	AGAGGGTCCT	CATTCACTGC	GGTGGACAGA	ACTCCTGACC	ACCTAGAATT
18201	TACCATTGTTA	TAAGATGTAG	AACAAAGCTG	GAAAAAGTAA	GCCTTGGGGA
18251	AATTGATTTT	GTAATAAATA	GAAAACCTGT	TTCTACTACC	CTATTAAACT
18301	TTTCTACTTT	CCTTCATTCT	CCCTAAATCA	TTTCCAATTT	GCCACAGACC
18351	ACAAATGACA	GAAAGTGACA	TTGTTCTCAC	ATCTTTGAAC	CACTGCTTTC
18401	CCAACCTCTC	ATTCACCTCT	TCTGCGAATT	TCTCTATATT	TTGTAGCCAA
18451	AGATTCTTGA	CATTTAAAAAT	TAGAGAAAGT	CAAAGTTGAT	GAAAAGTAAA
18501	TTTACTGGAA	ATAATCATCA	GTGAGAAAGG	AAAAGCCTGG	AACTGTATTT
18551	TACCTTGTTA	TCTCCTGTCA	AACAAAGTAT	CGGGAATCA	GACAAGAGTT
18601	CAGATCTTGG	TAAGATTAGC	CAAGTCTATT	CCTAACTTCC	TGTTTTACTC
18651	ACTGCTCATC	CGTCATTAA	TACGACTCTT	TAGGTTTCAG	CCGCCGGGTG
18701	TGGTAGCCAT	CTGTTTGTTA	GCAGCACCCA	GATAATTTCA	AAATGTAGAT
18751	TCCCAGATTT	ATCAAATCAG	AACTCCTGAG	GTGGGAGCAC	AGAAATCTAT
18801	TTAAAAAGCA	AACAAACAAC	TTACACACAG	ATCTGAGTAT	CATTTGTTTT
18851	TTGTTTTTTG	GACCACATTA	TCCTAAGAGT	GTCAATCCAAC	GTGATTTTCA
18901	AAATGTGACC	AGGAACCAAC	TGGGAAAAAA	AAATCACATT	TGGTAGTTTT
18951	TAAAGTATAG	AATTTTAAAC	TCACTGAATT	CCACTATATT	ATATGCTATG
19001	ACCTCATATA	TCTGTTTTCT	TTTTAACAAA	CTCCTCCAGA	TTATTCATAT
19051	ATGCACAGTA	CAGTTTGAGA	ATCAATGACC	TGGGGCAGAG	GTCTCCAAC
19101	CAGATGCCTT	CTAGAGGCCA	TGAAGGTAAT	GGAAATGTCC	AAAACAGTCC
19151	CAAATAATAC	AGTAGGGAGT	AGTGATATCG	TATGTCACCT	AAGAGTGCCT

FIGURE 3, page 6 of 27

19201 GGCCTTTCTA CAGCAGCCAG CCAGCTACTA CTCAGCTGCA ACCAGCTGTT
19251 ACTCACAGGG AACATTGCCA GATATTCTGA CTTTTCAAGG GAAGCCAGAC
19301 TGGATTTTTT TAATGTAAAA ATCCCCCTAA ATGTTGACAA CTTACTCACT
19351 TTTTAAAAAA CAAACTGCAT GCCTGCTAGT GCTGGAGGGT CACCAGTTCA
19401 ACACCTTCTG GACTAGGAAA TATCAAGGGA TTTGTAAAGC AGACAAGTAT
19451 TAGCCAGAAA CGCCTCACTG CCTGGCTGAG TAATGGAAGA TGGCATAGGA
19501 TAGGCCTGTA ATCATAACAA GTACAGTTCC TTTAAGGTAC AGCTAGAAAG
19551 AGCTAGAATA AGTATATAAT GTAAAGGACA GGTGGATACC CTCATGTGAA
19601 AGCAAGAGAC AAGAAAGAAA AAAGGCTCTA AAAGATAATG AATATAATCC
19651 ATTCTATTCA TAACATGCCC ATAAATGAAG TTTAGAAAAAC CTTACTCATA
19701 AAATAGAAAT AATGAGATAA TGATAACTAC TCATAGTGTT TGAGGTATTA
19751 ATTAAGTAGA ATAGCTTATA TCAAGTGGAT TAATGACAGC AACTGACATG
19801 TAATATGTAT TCAATTTTTT TAAATGTGAG TTCTTTATTT TCTGTTTCCT
19851 CAGGCTTCTG TTCCCTCACA CTAAAATTCA AGAGGCACTA AAGAAAAGCA
19901 ATCTCATGGC AAAAAGCTAA CACACTTCT TAATTTCAT GTGTGTGTTT
19951 AAAAAAAAC TGACATTTCT ATGTGATAAT TAACAGATT GTGATGACAA
20001 AGCCATTGAG TCCCTCATT GTCCTTTCCT CTAATTCTGC TCTTCTTCC
20051 ACTCTTTAGT GTTTCCAAAT TCCATGCGAA AAAAGTTGCT AAATAAATGG
20101 ACTTGAGAAT TCTTCTGGAT GATTTGGAAA AAGTGGATAA AGTCTGGGCT
20151 ACATTGCTCT AGAAAGATTG CTTTCATTTT ATTGCATTCT TGATATATCT
20201 ACTTTTTTAA AATATAATAA TTTGTATATA AAAACAATGT AATGGTGATG
20251 TTTAAAAATG TATTTACAAG GCAAAAGATT CACAGTTTCA CCACCTTAAC
20301 ATATATTTCC TTCCAATTTT TGTTTTTCTG GGAGGTTATT GTTTTCTGTT
20351 TTATTTGCCA TTGTAATTCA AGGGTCTATT AACTGTTTTT GCTCATAGTA
20401 ATCACTCAGA TATTTGTAA GGAATGAATG AATGAACTCC TGAAATTGTC
20451 ATGTACAACT GACTTTGTTT TCTACTTGCT CACTTTCATT ATGTCATGAA
20501 ACTTTTATGA GTCTCCACAG ACTTCAAATG AATCATATCT CTCTACCTGT
20551 TTTGCTGATT TTTTTTCTAA TGATAAAAAT CAAAAAATG TGAAACTTTA
20601 GAAAAGAAAT TTTAATCCCA CTATCCAGAG ACAATCACTC TTAATTGTGTT
20651 GAGATATTTT ATCCTAGCAT TTGTTACATA AAACTTTTTC TTTAAACATG
20701 GAGAATTTTG AACTCCTAGC CTTAAGATGC TCCCATTTTC AGCCTCCCAA
20751 AGTGCTAGGA TTATAGGCAT GAGCCACCAT GCCCAGCCTT AAACATGGAG
20801 AATTTCTATT TAGATTATTT CATTTAATAT GGCATTGCCA GTATTTCCCA
20851 TGTAATGGGA CGCCCATGGT CCACCCACAC CATTATTTTA TTAACGCTT
20901 ATTATTCCAG GCTCTTGAT ATGCATAGAA TGTTACATT TTCTTTTTTT
20951 TTTTTTTTTT TTTTACCATT TTTGTACAA AATGTCCCAT GATAAAGGTC
21001 TTTATTCACA GGCTCTTATT TCCTTATCAT ACATTCCCTGG AAGAGGAAAA
21051 TAATACTAGT TGATTGAAGA ATGTCCAAG ATCCTTTTCA TATTGAGTAT
21101 ATGAATGGCA TATTTTCCAC ATCCTTATGT ATTCCAGCCT GTCTTTGTTA
21151 TTTTCTTAA AAATAAAAGA CAGATTGACA GGGCATAAAA TTCTTTCCCA
21201 AATAAACTCA CTTAGATATC ACTTTATGAT CATTGGACAT TTAATGTAGC
21251 AGAACTGAGA AGCCCATGAT ATTTACATTT ATCCCTTTGA AGATCACCTT
21301 TTTAGTTTTT TTTCTTCCAA AATATTTGAC AGTTTTTCTA ATTAACAT
21351 TTTGTGAGAC ATGCTATGA ATAGTTCATT TGCATTGATT TTGTCAAGAT
21401 TAATTTGGGT CCATCAAATT GACATTTT TTAAATCCAG AAAAAAATTT
21451 ATCTTGATGT ATTTTTCCT AGCACTTCTG TTTTGTAAAT GGCTTCTTTC
21501 TTAGGGGCAT AATTCTTAGG TTGGGGTCT CTTGTTTATC CCCTATACAT
21551 CAGGTGGCCA AATAGTTTAT ATCCAACCAG AAAAATTTGA AAATGAAAAG
21601 GGAGAAAGTGC TAATAACTAC AGGGGACAAT AGGAATAAAA GAGGAGCTGT
21651 CTGAGTTCTA ACTGACACAC ACAGCCACTT TCCCTCTTT TCTCTACTA
21701 TTTATAGTTG TTGATCCTTT TGTTCTATAT CCTGTAAGAT TTTCTTAATT
21751 TTGTCTTCCA CTTCACTAAT TCAATGTTG CAATATTTGC CAAGCTTTTT
21801 ATGTTTCCAA AGCATAATTT GAAATCTATT ATTGGATTTA CTTTTCCTGT
21851 TTCTTTCCAT ATCACACTTT CTTTTCATTC CATCCTATTA TTGTCTATCT
21901 CAGCATGTC TCTAGTGAAT TTTTCATGTT TCATAAAAAT CAAATATATCC
21951 TGCTTGTTAT GTAAATGTCA GGGTTTTTCT TTTCCAATTT TCTTTTGGCT
22001 TCTAAAGTAA ATCGTTTTCA TAATGCGTCC TTTATTTTCT TTGCCTTGTT
22051 CTACAGCGTT TAGTTTCTTT GCCTTTTCT ATACCATTTG TAAGGGCCCA
22101 TGTTGCTTCC TTGTTTTGTT TTTAATTTT ATTTTTCAGA AACAGGATCT
22151 TACTCTGTTG CACAGACTGG AGTGCACTGC CACAGTCAGA ATTCACTGAA
22201 ACTCCTGGGC TCAAGGGCTA ATTTTTTGTT TTTACTTTTT ATTTTTGTTT
22251 TGTAAGAGATG GGGGCTCCAC TCTGTTGCC AGTCTGGTCT TGAACCTCTG
22301 GCATCAAGCA ATCCTCCAC CTCTGCCTCC CAAAGTGCTG GGATTACAGG
22351 TGTGAGGCAC CATACCCAGT CTTGGTTTTG GTTTAATTG GCGGAAGAA

FIGURE 3, page 7 of 27

22401	TTCTTATCT	AAAAACAATA	AGGAACCCAT	CCATACCCCTC	TGTTTACCAA
22451	CCAATAGGCT	ACTTGGACTG	CACCGTCTAA	CTTAGGATTG	ACTTATGTGC
22501	TAGCTAAATC	CCCTTCTAAT	AAATGCAGAT	AGACACTAAG	TTACCAAGCA
22551	CATTGCCTAA	ACCAAATCCC	AGTTGTCAGC	AGCCATCAGG	CATGTGGTGT
22601	GATTCTGGGT	AACCCAAGCT	GGAATTTATT	TATCCTCTGC	TGCAATGCTT
22651	TCTGATACAC	AGTTCCTCTA	CCTAGTCATC	TTTTACTGAA	TACATACCTA
22701	GGCTCTATCC	TAGACTTCCT	GGTTCAAAAT	CTTAGAAGGC	ATCAGCTTGT
22751	GTATGTAACA	AATTTCCAG	GTGTTTCTGA	TATGCCCTTT	TTGCTGGGAA
22801	TCACTGATCT	GAACCAATCA	AGTACATTTA	AAAATAACAC	AGCTGTCAGT
22851	TCCATGTCTG	CTGGGGTGAA	TGTCTAATTC	TTCTACCTCC	TATGGTGTTC
22901	TTCTGTGGG	AATTTAACTT	CTCCCAAACA	CGTTCTTCTT	GGAGATGCCT
22951	ATATTAAATA	ATAAAATGT	TCATCAGGGC	GGAGAAGTGA	TTGAATGTCA
23001	TTCTGACTGC	CCCAGAAAGC	AGAGCCCATG	CAGTAGAGGA	GCTCTAGTGA
23051	CTTTTCCAAT	TGAACTAAAT	GGCCGAAGGA	AAGGGATGGA	AGTCTTTAAA
23101	GGAATTAAGA	GCCAGAAGAC	CCAGATTACC	TGGGTTTGAA	TCCCATCTCT
23151	GCCAATTACT	AGTGGAAATG	AATGATTAAG	TCACCTAAGC	TCTCTATGCC
23201	TTAGTTTCCT	CATTCACAAA	GATAGGAATC	TTAATCTTAA	TCTTTTCTCA
23251	TCCCATAAGA	TAGAAATAAT	AGGATCTGTC	TCATTGAGTT	TTAATAAGGA
23301	TTAAATGAAA	TAATTTCTAT	GAAGTGTTAA	GAATTCCTTT	GTCCATAATA
23351	AGTGTTCCTA	GGAATATTAG	TTATATCATT	GTTACATGGA	ATAACGATCT
23401	CATATTTTTT	ATTCTGTTTC	CACTAGCTGG	TAGTTTGTGC	AATATCCTTC
23451	TGTTCTAGCA	ATAAGCTGAT	TATTAGTAGA	TGTTTTTAGA	GTGGTGAGCT
23501	TTCATATTTT	TGTGCTTAT	TTGATATTTT	ACTGGGAAGT	TGAGAGGCAC
23551	TTCATCAGAT	CAGTTCCTGC	ATTTTATTGG	AATCTTATGG	ATGAGTTCTA
23601	GAATGGTGAT	CCATCACTGT	AATTTGGGGT	TGAACAAGAA	GTCAGTCATT
23651	TCATTTCCAT	CCAGGCTTTC	CCACCATTTT	CTACTCACTG	CCTTGTCTAC
23701	CTCATTTGTT	CTTCCACTTA	GTTCTGTAAC	TTTGAAGCAG	CTCTGAAGTA
23751	CAGTGAACCC	CATGACCTGG	TTTGAAGCTA	GTGAAGTCCA	GGAAGAATTG
23801	CACTCTGTAG	TTCAAAAGGC	TCTTCTGGGT	GATAGTCATT	AAGAGAGAAA
23851	TTTAGTGCAA	AATCAAGATC	TTTCTAGGTT	TTCCAAGTAA	TTAATTA AAC
23901	CATCAGATAG	TAAGTGATT	GGTGAGACTT	AGTCAGTTAT	TTGAAGAGTG
23951	AAATTTTAAT	GAAAAGAATT	GTTAATTAAAG	TACAACTTG	TCAATTAGGT
24001	AGTTGGAAGG	ATAATAGGAG	AAC TTTATGA	TATCGTGAAT	TTAAATTCTC
24051	CAAGCAGTTG	CCATTCATTG	AGCTGAGGAA	AGAACAAGAG	ACTGGAAATA
24101	GAAATATTTA	GAGGCTTAGA	GGGGTAGCCC	CATAAAGCTG	AAATTCAGGC
24151	ATCTGAGGAA	AAGGGGTGTT	GCTCAGCTGG	TGATGGAATC	GTTGAGCTCC
24201	AAGGAGGGAA	CCTAAAGAGC	TCAGGTTTCA	ACCTTTACCC	TGCTGGTTGA
24251	TGGTATCTCC	GAGGGCACGG	AATAAAGTGG	TTCTACAAAT	ATTGAAAAAT
24301	AGCTACTGTC	TCTCATTAATA	GATTTTTTGA	AAGTCGAGAC	TTGGAGCCTT
24351	AATTATCTTA	GCTCTCTAGT	TTCCATTAC	CCACCCCTAC	TCTGTCCAGC
24401	TTTCCAATGC	TCATTTTCGCT	ATCAATTATC	TCCAGTTTTG	AAAGACAGCA
24451	CTAGATTCCG	CCACTGCCCT	GGAAAAGAAC	CACTAACAAG	TTGAAGAAGT
24501	ATTGCTGGGG	TGATGCTTAC	AGGAACCTAG	AGCAGTCAGG	TAACCCACAG
24551	AAAGCAAATC	AAAGACAGAA	AAGACAGACA	AGAAGCAGCA	AACCCTTTCT
24601	CTTCTTCAG	TCTTGCACTC	TCCCCCTAGT	GGCCTCTACT	GGCAAAGCCC
24651	ACCAGAACCA	GAAACGTAAT	GGGTGGAATA	CAGTAGTCTC	CCCTCTTCAA
24701	GGTTTTGCTT	TCCAAGGTTT	CAGTTACCCA	TAGTCAACCG	AAAATAAACA
24751	ATTCTAAGT	TTTCCATTGG	ACGCTATTCT	GAGTATCATG	ATGAAATCTC
24801	CCACCGTCCC	ACCCTGACCT	ATCGGGATGT	GTGAATCATC	CCTTCGGTCA
24851	GTGAATCCGG	GCTGTATGCA	CTACCCGCCT	GTTGGTTATC	AGCATCATCT
24901	GCCCCTGACA	TCCAACCATA	GATGTCATCA	TGGCTCGATG	ATCCAGGATC
24951	TCCCAAAGCA	GATGGTCCTC	CTGACATATG	GTCAGAAGTT	CAGTAGTAGC
25001	CTAATGCTCT	GTGACAATGC	CAATGCCATT	CACCTCACTT	CATCTCATCA
25051	CATAGGGATC	TTATCATCTT	ACATTATCAC	AAGAAGGACG	AGGGTGAGTA
25101	CAGTACAGTA	AGATATTTTA	AGAGAGAGAG	GCCACATTCA	CAAACTTTT
25151	ATTACAGTAT	ATTGTTATAT	TGTCCTATTT	TGTTATTAGT	TATTGTTATT
25201	AATCTCTTCA	ATTAATTTAT	AAATTGAATT	TTATTATATG	TATGTATATG
25251	AGAAAACATA	GCATAACTGC	TCAGAGCTGT	GGAGACTGCA	ATTTTCCAAT
25301	TCATCTGAAC	TAAACAGTCA	AAATCAGACA	ATTCAAACCT	TAGAAGAAAT
25351	CTATTAATGT	GAATGATTG	CTTTAGCCAT	CTGCTTTGTC	ATTCAGAGTG
25401	CCTAAAACCA	TGTCTGGTTG	CCAGTCAACT	CTGTACCACA	TAATGTTACC
25451	ATAATTGTTT	GGCAAAATTT	CTACCTGGGA	TACCTCATCA	CCAGAGGTTG
25501	ACTTTTCTCA	GTCGTCCCTC	CCTTAGGAAA	GCAACTTGAG	TAACTGATT
25551	ATATAATTTT	GCCATAAATA	GAAATCCCTG	AGTTACAACA	TGAATCCACA

FIGURE 3, page 8 of 27

25601	GAACAGGGTG	GAAATTGGAA	GGTATCAGCC	TATTGCAGAT	CTCTAAACTA
25651	GAGGAGATCT	CCATCCAAAC	ACTAAATTCT	TACCCTTGAA	TATGGACTTG
25701	GGGCATCAGT	CATCTGCCAA	ATCTTAATAT	ATACATTGAT	GCCACTGATG
25751	TCTGCAAAAG	ATTACCAGTA	TTTGCATATA	CTTGACATTC	TCTATACAAC
25801	AAGTGAATTT	TTATGTAAAT	TTGTATGTAT	GCATGAGACT	CAGTAAGAAT
25851	TCTCCAAAAT	AGAAAGATTT	GTTGGAATAC	CAAAATAGCTA	CTATCTGTCA
25901	TTTAAGATTT	TTTGATAGTT	GAGACTTAGA	GCCTTAATTA	TCTTAGCTCC
25951	CCAGTTTCCA	CCCACCATC	TCTACCCTGC	CTAGACTCCC	AGTGATCATT
26001	TACCTATAAT	TACCTCTCCT	GTTCTTAAAG	ACTGAACTGA	TACTGTCCAT
26051	ACTTCAAAATC	TAATCTCTGT	ATTTTCTTCT	CCCGGACCCC	AAAATCACTG
26101	AGATTCACCC	GGAGTTCTCT	AAAACAGGAT	TTCCAGGAAA	AACAATTTAG
26151	AGAATAACAG	TAAAAGAGAC	CAATTTTATG	TAGAATAGGT	TTACGTTCAA
26201	GGCATCCAAG	CAACACTTTT	TGAAATGTTT	TTTAAGCCAT	CATGTTGATA
26251	GATCATAAAA	TGACATCTAT	CATTCTCTGA	GACTTTTATA	ACTGAAAAAG
26301	GAATAAATGC	AGTGTAGAGT	CAGGCTAGAG	TGTTTCACTT	CCTTGGGGCC
26351	TTGGGTACTT	GTATAATAAT	TTTTAAAAACA	TTTTTGTACT	TGTGTAATAA
26401	AAATAATCAC	TCACCTTGAT	ATGCATTTTA	CAATTTGCAA	AGTAATTTCA
26451	GTCACTTAAC	CTTGCAATTA	CGTAAAAACC	TAAAACACT	TTGAGAGGAA
26501	GGTATTATTTC	TCCCAAATTA	CAAAATGAACA	AACCTGAGCGT	TGGGTAATTT
26551	ATTTTACCGA	GCAAGTAAAA	AAATAAAATT	TTCTGATTTT	AAGTCCAGTT
26601	CTCTCTCCTC	TAAATCACTA	CAGATGCAGA	GGTCCTTCTG	AGAACTTAGA
26651	CGGCAGCGTG	AGCTGCTACA	ACATCAACTA	TGGAATTCGT	AGGTCCTAAC
26701	TTCCCTCCTG	ACACATTAAT	AACCAGGCTC	TGCTGCCTCC	ACAAAACCAA
26751	GTGTATTCTA	CCAAAGGTCC	CATAAGCAGA	AAATTGTACT	CTGTTTCAAT
26801	AAATGGTATA	TTTTTTAAAG	CTGCCTTTAG	ATTACCCCTT	TAGCACCTTG
26851	AAACTGTATT	TATTATCATC	TGAAGCTGGT	GACATAGATA	AATAATGAAT
26901	CTTATTCTTT	ACCAGAAAAG	GTCATTTGAA	TTTTCTGAGA	CCTATTTGAC
26951	CTCAAATACA	CATTAAACATA	TTTATCATT	GCTTCCTTTA	TCATGTCCGG
27001	CCTCTAGAAA	TGGGTAAGCA	TCTCATCTTC	CTAGAAAAAT	TCAATTTCAA
27051	AAGAGAAGAA	AAAAAAAACA	AAGAGTTAGA	ATACAGGTTA	TGGCTAGGAA
27101	ATAGTGAGCA	GGCTGTTTAA	AGAGTGAGTC	CATTGCCAAG	GGTCATAGGA
27151	ATATTTTGAA	ATTGCCTGTG	TGTTACTATC	ATTTAGAATC	CTTTCCAAAG
27201	GTTTCTGAAA	ACATTTACAA	GAGTTAAAGA	TTCAATCTTG	AGCTTTCTAC
27251	TATTGTGTGG	GATTTATAAA	ATATGTCCTA	TGACATATTC	ATATGTTGGA
27301	GGTTTACTGC	GAAATTTTAT	GTGACAGTCT	GCAAAGTTAC	TTTGAGGACT
27351	TTTGATAAAC	ATCTGGGAGA	TGTTAGCATA	GACCTTATAA	TGTGAAAGGT
27401	AGATGCTCAC	TCATCTAGCA	TAAAAATGTC	AGGCTAGCCA	TAGAAACGCA
27451	TAAGACAAAT	CACACTCCA	TTATTTCTGA	AGATTTATGC	TTTGAACAAA
27501	GATGATTAAT	TTAATTCTGG	ATCTCTTAGC	CAAAATGACC	CATTACCATA
27551	CTTCTATTAT	TTCTATAATA	TAAAACAGGA	ACCCCTATCA	TGGGTAGGGG
27601	GATCATATAC	CGTATTTTCC	TAGACCACAT	TTAATATNNN	NNNNNNNNNN
27651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNCTG	GAGTCCAGGC
27701	CCAAGCCCCG	GGCCGGCCAT	CTTTTATTTT	TAAAAGACCC	TTTTTCATAC
27751	CTCCCTCCTC	AGGACAGCAG	TAGACCCTTA	GCATGTAATC	AAGTTCTTGA
27801	AGGGGCTTCT	AATTCCTTCT	CCCCTGTTGA	ATGCTTCCCC	AGTGATGATT
27851	TTTCCCAACG	CTCTGTTTTT	TATCATTACA	AGCTTTCCCT	CAAAAGCTTA
27901	ATTTGGAAAG	ACAAAAAGAC	AAATTCTCAG	AACAAATTTT	AAGAATAAAT
27951	TTTAAGACTC	CATCCTACCT	ACCATCATGA	GATGCTACAG	TCATAGTTTT
28001	CCTCCTGTAA	ACCTCTGTGG	AGTCCTCACT	TTTTATTCTA	CATGTATCAA
28051	AGTCATTTAA	ATCCTTCAGA	GAGTTTAAAG	TGCAGCATAA	ATTCTATTTT
28101	CTATTACACA	TTTTTTGGTA	CGTAGGGACA	GGCCACACTT	CTTTGTTTCA
28151	TACAAGGGCT	TGAGATTTTA	CTGAGAAAGG	CTTCCTCATT	TCTTAATGCA
28201	TATTATTGAA	TGTTTCAGCA	GCTCTTAAAA	GCAATCCAAT	AGTTCCCAAA
28251	ATAAACATAG	TTAATTCCAT	CAATATTTAT	TGTGCCCTTA	ACGTATATAA
28301	GACATTGTGC	TAAACATTGT	CAGGATGTCA	GATGCTTGCT	CTTGAGGACA
28351	ACAATTAACA	AATGTATTCA	TTGGAAGACT	ATTTCCACTT	CAATTATATG
28401	ACTGCTAATT	TGTGACTTTT	CAAATAAGTT	TCTCTTCTTC	CATGAAGTTT
28451	GTGAATTCCA	CAAGTAGATA	AAATTGTGAT	ACAAGTTACA	TAAGTGTGTT
28501	TATAGGCACA	GATTTCACTT	TTACACAGCA	GAATCCAGGT	TCAGAAGATA
28551	CAGAAAAGTA	ATCAAGCCTT	TAACATTGCA	CTACACACTT	AGTCTTGATG
28601	TATGTATTGA	AAATTATCTT	TTACTTATTA	GAAGTGTCTT	CCTCAAGGGG
28651	CTGGAAGTTT	AGGAAACTAT	GATGCATCCA	TCTTTCACAT	CATCCTTATA
28701	AAAATACCTG	CATTTTGCTA	AGATTTCTCT	CCATTAATTT	TAAAAAGAAA
28751	CAAAAGTAAT	TTCTTCTCCT	TATTGCGTAT	GAGATCAAAG	TTTAACAAAT

FIGURE 3, page 9 of 27

28801 GAGGTCTTAA TAGCGATACC AAGAAATGGG AAGCCATAAA TGAGACTGCC
28851 TATATGGCAG TAGACAAGCT TGACAAACT CCTCAACCAA ATGTATGATT
28901 GTGTTACTTC TGATATTCAC ACCAAGAAC ATCCACCCTC TGGTACTCTT
28951 AGCAAAAATT ATTAGGAAAT CAGTTGTTAG GAATCAATAG TTCCATTAGA
29001 CAGGAAGCAT AGTTTTCCAA ACTATGGGAA TTTTATCCCA GAACTATGTA
29051 TCACAGTGAA ATTAAAGGAT TAAGCCTCAT AAGAAAGCAA AAGTACCCTA
29101 TGTAAAGTC TTTGGCCAAT GCTTCTAAT TCTTCTTTT CATATCTTTA
29151 AATACAGAAT CCCCTTCACC TACAACTGAG TTTAGAAAAT TCATTAAGTT
29201 CTGATGCTGA TGTCACGTGC TCAATCCTGA CCATGAACAA CTGGTACAAT
29251 TTTAGCTTGT TGCTGTGCCA GGAAGACTGG AACATCACC GACTTCCTCCT
29301 CCTTACCCAG AATAATTCCA AGTTCCACCT TGGTTCTATC ATCAACATCA
29351 CCGCTAACCT CCCCTCCACC CAGGACCTCT TGAGCTTCTT ACAGATCCAG
29401 CTTGAGAGTA TTAAGAACAG CACACCCACA GTGGTGATGT TTGGCTGCGA
29451 CATGGAAAGT ATCCGGCGGA TTTTCGAAAT TACAACCCAG TTTGGGGTCA
29501 TGCCCCCTGA ACTTCGTGG GTGCTGGGAG ATTCCAGAA TATGGAGGAA
29551 CTGAGGACAG AGGGTCTGCC CTTAGGACTC ATTGCTCATG GAAAAACAAC
29601 ACAGTCTGTC TTTGAGCACT ACGTACAAGA TGCTATGGAG CTGGTCGCAA
29651 GAGCTGTAGC CACAGCCACC ATGATCCAAC CAGAACTTGC TCTATTCCC
29701 AGCACGATGA ACTGCATGGA GGTGGAACCT ACAAATCTCA CTTCAGGACA
29751 ATATTTATCA AGGTAGGATG CAAGGTCTCG GTTATATCCC CATTATAGG
29801 GCCATGACAG AGAGTAAAT TCCCCTATCT GTCCGCTTTG CAGAAATCTT
29851 GACTCTGAGT AGCTTTAAAC TTTAATAATA TTTCTTAGAG GATTCTGGT
29901 TATATAGGCT AGTATTTCAT GATCTGCTAT CTGTAATTTG ATCTATAAAC
29951 TTGTAAGTAC ATGGTATAGT GGGAGTGCTC AATCCTGCCT TTAACCTTG
30001 GTTTAGCTTC TTAGTAGCTG TTGTGTCTT GGAAAGTTAT TTAAGGCTC
30051 CAAGCCTCAG TTTTCCAAC AGCAAAATAG AATAATGAAT AGCTTGGTAT
30101 AGTATAATAA AGATACCATG AAATTTATAT ATGAAAAGTA CCTAATACCA
30151 TGCTTAGCTT ATAGTAGATG CAAAATAAAT GTTTCTTTT CTACCCACTC
30201 TTTTCCATAT CAATAAAAT AATCAAGTTT CTCTAAATCT ATACAAAGAA
30251 AAAATTAGTC AAGCAAGAAA TGGACTTTT TCCCTCCTCC CTGGCCTTGA
30301 TGCTTAAGAC AGTATAGAGT AGTAAAGGCA AAGACTCTTG AACTCAAGTT
30351 CAGACCCTCT GAACCTGAAG TTTAGCACTG CCACTTACTA CTTGAGTGAC
30401 TTAATCAAGT TACTTAACTT CTCTGAGCCT TAATTTCTT TGTCTATAT
30451 TCATTTATGT AAAATGGATA TAAGAGTAGA TCCTATTACC CATAGAATCA
30501 TTGTGATTGA TGATTGATAG ATAGATAGAG ATAATAGATG ATAGATTGGT
30551 AGTAGACAAA CAGGATACAT TAATAGAACT AAGAGTTCCA TAAGAGTATT
30601 ATATATATAT AAAATATTAT TTATTTGTTT ATACATTCAT AATTATATTT
30651 GTTTATTTT AAATTAAATA CATTTTCTGC TCCGGTTCCT CAATATGATT
30701 CAGAACTAG AACGAAAATG TCCATTTAAA ATAGAGAACA ACACATCATA
30751 TGGAAATGGT TTTGGTGATT CCCGGGAAAG GGGGAATATC CCTTAAAGGA
30801 TATTTTATTA GGGCTTAGAC TTTCTTCTGA AAAAGGACCA CCTGTAGTCA
30851 GAGAGGCCAA GTCAGAAGAT TATTATTCT TCAAAGACGA ATGTTCCCT
30901 GTAGACTAGG CCCTGTTTT AGGCCATCCT GGAACAGTGG TATCTGACTA
30951 TGTTGAGGAC TACAAGGCAA ACTCATAACT TCTTACCTTT AAAAAAGACA
31001 TGATAATGGG CAACACAGCG AGACTCCGTC TCAAAAAAAA AAAAAA
31051 AAGACATGAT AAAAGGTCAA GGGGTGCAA TAGTTGTACA TTTAATTTTA
31101 CATATATATA TATATGTATG TATAAATTGA GCACCTACTA TGTACAAGGC
31151 AATATGCCAA ATGCCATATG TAAGGGAAAA GTGAAAGACT GAACACAACC
31201 TGTAAACTCC TTAAAGAATG TTGTTAATAA AATTTTCAA ATATATTTAC
31251 TACAAATCTA TTAGTAATAA AATTAGATGT TCTATCATCC TCTGAACTTT
31301 CCCTTTTTC CATATTATAA TTTCCATAAG ATTAAAATCC ATGCATATTT
31351 TATTTTATAG CCACTCTCCA GATTTAATCT ACTGTTGGCA AGCTCGCACA
31401 TAATTAAGGT TCAGAATTTT ATCTAAGACA AGAAACATTC TCCTTTACAA
31451 CAAAAATACA AGCAAAGTTT TGATTTATAA TTTCAAATAG TCATTGTTTT
31501 GGAAGGACAG TCATAAACAG CAGCCAGGAA AAACCACTTA TGAAACTAC
31551 ATTGAGTTCC TTAGCATCTT TTTGTCTCAT GTAAAAAGGA GAATGCAAGA
31601 AAAGTGATTC TGTTTGAATC CTAAAAACGT TTAGAACTA CAGAGAAGAA
31651 TATTTGTTGA CTTAAGTTGT ATATACCTTA GGGTCTCATT TTACCAAGAT
31701 CAGACTGATC TTTCTGGCTC CTCAAGATTT AATTATATT AAAATTATGT
31751 TCTTCCTTCC ACAAGGATCC CATGGCATTT TGTATATAAA ATTATATAGT
31801 GGTACTTGGC CCATTCTATC CTCATTATTT GTCTGTGAAT CTGAGTTCTG
31851 AGCTAGAATA TAAATTCTGT TAAGTGAGGA CGCTATTATT GAGCCTTATC
31901 CAATGCCTAG CCTATAGTAG GCACTTAATA GGTATTTATT GAAATTGACA
31951 TAGAGGCCAG GGGCGGTGGC TCACGCCTGT AATCCCAGCA CTTTGGGAGG

FIGURE 3, page 10 of 27

32001	CCCGAGGTGG	GCGGATCACG	AGGTCAAGAG	ATCAGACCAT	CCTGGCCAAC
32051	ATGGTGAAAC	CCCGTCTGCT	CAAAAAAAAA	AAAAAAAAAA	AATTAGCTAA
32101	GCATGCTGGC	ACCGCGACTG	TAGTCCCAGC	TACGCGGGAG	GCTGATGCAG
32151	GAGAATCGCT	TGAACCTGGG	AGGCAGAGGT	TGCAGTGAGC	CGAGATCGTG
32201	CCACTGCACT	CCAGCCTGGC	GAGACTCCGT	CACAAAAAAA	AAAAGAAAGA
32251	AAGAAAGAAA	GAAAGAAAAA	GACAGGAAGA	AGGAAAGGAA	GGGGAGGGGA
32301	AGGGAAGGGG	AGGGCAGGGG	ATTGACATAG	AAAGAAAGAA	AAGTAACTTT
32351	CTGTTTTATT	TACATCCTAC	ATTATCTGTT	GCTGTAGAAG	AAATGGATAG
32401	ATGGTAGATA	TTGTCTAAAT	TAAGTACTTT	TTAAATTTAC	ATAAATACTG
32451	GATGATTTCT	GTTTGGTTTT	CTTTCTCTCT	CGCTCTCTCT	CTCTCTCTCT
32501	CTCTCTCTCT	CTCACTGTCT	CTCTCTCTCT	CTCCTTATGG	TGGGATGTTA
32551	GGTTAGTATA	TAACTCTGCT	CCTCGTGGCT	ATTTATTCCA	CAAACTTTGA
32601	AGCTGTAAAA	GAAGATTGTG	AGGTTTGAAG	CCCAGCAGTA	ATTTACACCG
32651	GGTCAGTGAC	AATATTTTCA	TCATGATGAA	TGTGTTTGAG	AAATGAACCC
32701	AAAGAACTTC	AGGAAGTATA	CCTGAGACTT	TTTAAACTCC	TGAGGGATAC
32751	AGGAAAAGAG	AAAGTTTGAA	AAGTATTCAG	GAACAAGTCA	AGGGAATGA
32801	GCAAAACCCA	GGAAGGAGAC	TTTATAATGA	ATAACTGAAA	AGCTGCCTTA
32851	AGCCATAAAA	CATCTGTGGA	TTTTCCATCA	TCCTTATTTA	TTTTGTTTAA
32901	TACAGTCTTT	CAGAAAGTAA	GTTATTTCTCA	GTCCATAAGT	CACATCTGCA
32951	TACAGTAATT	TTCTTAATTG	TTCTTAAATT	TTGTAAGGCT	GCCCCACTT
33001	CTTCACTGCA	TAATGAAAGT	CGGGAGGATA	ATGAGCCATG	AATAGTGGAT
33051	GTCAGAGTTA	CGCAAGTTTT	CATTTCTCAC	ACAGTCATTT	TCAGTTTGGG
33101	CTGACAGAAC	TGTTAACATC	TTAAAAATGT	AATGAAATCA	CCAAAAACAG
33151	GGCATTTTCA	GCTAGGCTTT	CAGATTAGAA	AAGTCATTTT	TCATGGCAGA
33201	CTACACACAC	ATAATTACAG	GTATTAGAGA	TTTTATTCTT	CCTAGGTCCC
33251	CACATGCCAG	AGCAAATGTC	CATAATAACT	AAATGTAGAC	AAAACATTCA
33301	GGGACCAAGT	TCATAGCATG	ATCTTCAACA	ATCTTCAACA	ATATATTTAC
33351	AAGTTTGTG	TTGTTTGTG	TTTGTTTTGT	AGACGGAGTC	TTTCTCTGTC
33401	GCCCAGGCTG	GAGTGCGGTG	GCGCAATCTC	GGCTCACTGC	AAGCTCCGCC
33451	TCCCTGGTTC	ACGCCATTCT	CCTGCCTCAG	CCTCCCAGAT	AGCTGGGACT
33501	ACAGGCGCCC	GCCACCACGC	CTGGCTAATT	TTTTGTATTT	TTAGTAGAGA
33551	CGGGGTTCGA	CCGTGTTAGC	CAGGATGGTC	TCGATCTCCT	GACCTCATGA
33601	TCCGCCCGCC	TCGGCCTCCC	GAAGTGCTGG	GATTACAGGC	GTGAGCCACC
33651	ACGCCAGCT	ACAAGTTGTT	TTTTTAAATG	TTAGTTAATT	GGAGCAATTA
33701	TTGGTGAAT	ATTATTTTGA	GAATACCTTT	ATAAGCGCAT	TTGAATGGAT
33751	GTTTTTGCC	AGCCAATGAC	CATGTGTTGA	ACTATGGTCG	GGTAGACAAA
33801	ATGAAGACTT	GAATTCTGAC	ATTTAGGAGC	TGACAGTCTA	GTGAATGAGG
33851	TCGAAATGTA	AGTAAATGGT	TATAAAACAA	TGGGATGTGT	GATTTATAAG
33901	AATAGGGATA	TATTCAGATA	TACAGAGGGA	AATAATGAAC	TCTTCTCAGA
33951	TATTTGGTGA	TGAGGGAAAA	ACACTGGACT	TGGATCCTGT	AAATGAAAAG
34001	GGGTTTTAGG	TTTGAACAA	ATTCACAGAC	AAAGGAAATG	ACATATGCAA
34051	ATTACCAAG	ACCTGAAAGA	TAAATTTAGT	ATGGCTGGAG	CATTAATTGC
34101	ATATGAAAAG	GATTTGAGAT	GATTCTGGAA	AAGTAGTCAA	CTTGATAACA
34151	TACCACACAG	AGTTTGTATG	CCAAGCTAAG	GCATTTGCAT	TTTATATGCA
34201	GAGCGGATGT	CAACTGAATG	GCAGAACATT	AGCCCTGACA	TATTTATAAA
34251	AATCATAATC	CTGTAACAAA	TTAACCAAAT	AAAGTAATAC	AGTATAAAAG
34301	CTTTGCAAGT	AATTTTTTTT	AAACATTAGG	ATATAAACAT	TGTTTTTATT
34351	TCTCAAAATT	GCTTTTAGCN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
34401	NNNNNNNNNN	NNNNNNNNNG	GATTTATTTG	CCCATGCTTG	CCTAAAGGCC
34451	CTAAGTGCTA	ATAAATTCAT	TGTTCTTGCA	GAGAAAGCTC	CATGGGTTGG
34501	TCATAATCTG	GTCAATCAAC	CTTATCTTTT	CTTTAAGAAT	CCCAAGAGTC
34551	TGGACACTGA	CAGACACATT	GTTGTCTTTG	GCCTAAGCAA	GTCCCTAACC
34601	CTTTGACTAA	TTAGCCATAG	TTCTGATCGT	TTATTCTAGC	TTCTAATTGG
34651	GATGTAGAAT	CCTCCTTTTC	TTCTTTGAAC	AGCTTCTGCA	GTTCTGTTCT
34701	GTATCCCTTG	AGAAATGAAT	CAAGTAGTCA	GTGATTAATT	CGTTCCTTTG
34751	TTACATGTA	TTTGCTGAAA	TTAACTAATA	TAAGGTCAAG	CCCTTGCTCG
34801	GAGACAGAAG	GAATAGAAGG	TAAGAGGCCG	TGCTGTCAGG	AAGCTAGCAA
34851	AGCCCTTCTG	TTCTTCCACC	AGGAGAAGAC	ATGCTCATCT	AATTATTCAT
34901	AGTTTATATC	ATTTTCTTTG	ACCTGGCTGT	TTGGTTATTA	TTCTTCACTG
34951	GGATAACACC	CTTTATACCA	GTGATTATAC	TCTGCTGGTG	CTAGCATTTT
35001	TAGGCCCTCC	TCCTCAGTTG	ACCTGAATTT	ATGATTCAAC	ACACTAGAGA
35051	CTCTTTTCTC	TTCAAACAT	GTTTATCTTT	ATGCTTACAC	AGTCAAGCAA
35101	TATCAATACT	CACTGGCCTT	TTATTAATAA	TAAATGTTTT	CATTATTGAT
35151	TCTGCATATA	TTCTTAAGAA	CCTATTATGT	GCCAGGCATT	GTGAAGAATT

FIGURE 3, page 11 of 27

35201 AGAAAATCTT CAAAAAATG AATCAAACAA ATGTCCATTC AAAAAACCCA
35251 ACATATAAAT TGATAATCAT CACAATTTAT TCCTGAACAT TAGCAGAGTT
35301 CCAGAAGAAC TTGAGGCTGA ATGCTATAAT TCTAATTTTG GATATTCTGT
35351 CCACGTGCAC CATAACTTAG TGGCATTGTG ACCATGTTAA AGTTATTTAA
35401 ATGTCATAAA CTTTAATTTC CTTAACTATA AGGATACCTG TCTCAAAAAG
35451 TTAAGTAGTA TGTCTAAGGT CATTCAAGCTA CTACTTTGCA GGGGAAGATT
35501 TAAACCAGTT TATCTAAATT CATATTCCAG GCTCTTTCAG AATTAGTATA
35551 AGCCTTCCTT AATTGGGATC CTATTACTTT TTAATTTATA TTCTTTCAGC
35601 CTGAAGCTAT ACAAATCGCA TCTGCATACA GTAATTTTCT TAATTGTTCT
35651 TAAATTTTAT AAGCATTCTT GTCCTAAGGA CCTCTACCAA CACAAACTGG
35701 TTAACCCACG TATTTCACAA TGTACTTAAA AGAAATGCAG TTGCATTAAA
35751 CATGGAAGCC AGGGGTGGGA GGCTGCTTAG CACTAGCTCC CTGGAGTCCC
35801 GAGAAGAACA CATTGCTTAT GGCTGATCCA GTATACCTAA CTCTTACTCC
35851 TAGGTAACT TTCTCTGCTA GGGCTACTAA GGTGTTGATA CTTCTAGAAA
35901 AGACATGTTT GGGTTCATGA ATCTCAGGGA TCAACACTTA AGGTCTGTGT
35951 GTTCAGATGT TTTCAGTAGG AACCCTTGTC AGCCGAATGA TCTGCTGGA
36001 ATTCTTTGAA ATTACCTCTA CTCCAGGTCA CTTAAGTCAT GCCAAGAGAT
36051 GAGTCTAAAA TTTTCTCTAA GTCACTGCGG GCGGGGTCG GGGGGACTCA
36101 CATAAGGTAC ACTGGAAAAT GTATATTCCC TCAAGACTCT ATTTTGATGA
36151 ATACAGCTCA AATTACTTAA ATCTAGGATT CAGCAGATTT TAACTGTGG
36201 AATATTTCCC AATTAGGAGA GCTTCCAAGC TTTTATGTGC CCGAGAAGGA
36251 ACTGTATCTT TGGTTGACTT TTTCACTTTA TATGCATCTA CTGTAAAATC
36301 TGGAAATCTG CCAAAAAGTA TGAAACTATG CAGAGTAATA CTGAAGCTCT
36351 ACTCTGATTT TCAGATTTAT CTCTCAAGAC CACTTCAATC TGCATACCCT
36401 ACAACCCTTA GGATAGATGT TTGACTGGTA AATACTGCAT AATGTCTTAT
36451 TGCCAGGGCT ATGCCAAGGC AATACTTGAA GGGACATCAA CACCTTGGCT
36501 AGCACTGGGG CCAGATCCAA AGAGCCAGAA TGAATTGAGT TGTGATGTTT
36551 CACAGTGTGA CCAGGAACAG TTCAAGGGCT AGGCAGAGAT CATAATTGTA
36601 TAGAAGAGTC AAGGTTTATA CCAAAAAGCA AAAGGTAAGG ACAATGATAT
36651 AGGGGTATGG AATAAATCAA ACTGTTTTAG GTAACATATC AGAGAAGAAA
36701 ACACAGAGAG AAAAAAGTGC CAGCTTGCCA GTTACAACAC ATGGAAGCAA
36751 GAAAAAGAGA TTACCATGTC TGTGTGTATG ACTTCTTGAT GTAGCAAGCC
36801 CCTCATACAC ACCAATAGTA ACACAGCACA AGACATGTAT TAAATTATGA
36851 CCCCACCACA TAGATTAGTA TTGTTCTCTT CTTGGTGAAA ACTTCAAGAA
36901 AGGTAGGAGC TCTCCTTCCA CCCTACAGTT TCACCTATAA GTAATTGAAT
36951 TTTGCAGATA TTGACAAAAC ATAGACCCAA ATGATTCATA TATGTGTACA
37001 TATATGTTTA ATATATTACT AAATTGCTGT TGACCATTAA CTGATAGAAA
37051 TATTTTTTAA AAGATGAGCC TTGTATGCAA TTTTAAAAGA TGACATAATC
37101 AGGGATTATA GCGTGAGGG CCTTCTGTTT TGAGGATGGA AATTAGCAA
37151 TTTCCAAACC CTATTAAACT CCTTCTCATC AGAGAGGTTT CCCATTGAAC
37201 TAACTTCAAT TTTTATACT CTCCATTATT CAGGAGGAAA ATGTATTGAA
37251 AGTTTAAATCC TTCAAACAAT TTGCAAATTA CAAATGCAAA TGTTTCCTGA
37301 CTTAATGAGC CCCATTCTTC TGGCAAAGTG ATGAAGATCA CTGTAAGAGA
37351 TACATGTCTG ATTTGAGCAG AAGACATGTG TTTAATTGCA TTTACCCCAA
37401 AACATTACTG AGCAGTTACC CTGGCCAAGC ACTGTGTTGT GCCTAAGAGT
37451 CAAAGATGAC TCAGTGAGAT AGGTACTTAG ATAAATGCTG AAGCCTCTTT
37501 CAGCTTGACA GCCTGCAGTG TTATGAAGCA GAGCAGTGGG GAAGGGAGAA
37551 CAGCAATTCT GTGGAAGATG CTGCTCTCCA AATCTGGAGT AGGTCCAAAC
37601 CCATGCCTTT GAGGGCCTTA TAGTCTAAAA AGTCAAAGAT GAGATAAATA
37651 AACTGCCAAA TTCTTCTACT TTAGGATAGT AAGAGGAGAG TCAAGGAGTC
37701 ATTTACATGT TAAGCTCAAG AAATCACGCA CATTTAATGT TTAATTGGA
37751 GAACGTGCCC ATATGTGGAG AAGAAAATCA AACAGAATTG GACCACAGGT
37801 AAGCTCTGTG GCTAAAATGG ACAAATTCAT GTTATCATAA AAGGAAGGCA
37851 GATACCACAG GGCTCGGCTT TGGTGAAACA AGCCACAAAA TGAAAGCTGA
37901 ACTAGTAACA ACTCGCCATC AAGTACAGAA AGGTTCCCTA GGGCCGTAAG
37951 AAAAGGGAAA AATGGTAAAA GAGACATAAA AAAATAAAGG GAAGTAAATA
38001 GATGGATCTC AGAAGGCAGT GGGGAAGGAG CTGGAATGGC GACAAATAGT
38051 AATTAACATA ACGATGGTCA AAGAGCTGCT TTAAGACAAG ATCTCCCACT
38101 AACAAGACAG AATATTGCA TTTCTGCTAT ACAAAAACCA CTAAAAAGAA
38151 AAGGGGGAGA GGGAGGGAGA GAGAAAAACA AAAAATTCAG AAAAAATAAA
38201 ATAAAAATAA GAGTATTAAA GAAACAATGA GCCAACTGTA GATTAACCCA
38251 CATGCCAGAA ACAGTGAGAA TACTGGAGGG AAGGGAGCCA GAGGAAGAGG
38301 TGTAATAATGA GAATAATTTA GGAAATCAGA GAGTTTAGGG GAAAGCCCTG
38351 AAAAAATAAG AAATATACAC ATGGAAAAAT ACAAATGTAA ACTATGTATA

FIGURE 3, page 12 of 27

38401	GTAATAATT	CAGATAACTG	GAGGTCTATG	GACATTGTGA	AATATCAAAA
38451	TTGGCTGTAA	GAGTTCGTAA	AGACAATCCA	AAGAGAGAAT	AGCTTAGGGC
38501	TCTTGAATGA	AAAAGAGCAG	AAAAAAAAAA	AAGACTACGT	AAGTGTGAAC
38551	TTGTGACAAA	TGCAAAAGTG	TAGAACTCTG	GAGAATGTGA	GTTTTTAATT
38601	AGAAGATTCG	TCGTAGATAT	GAATCACATT	AAGAAAAGAT	AGGATTACTG
38651	AACATCTATG	TCAAGTTTCT	CTTTCTCACA	GAAAAAAAAA	AAAGGAAAAG
38701	GGGAAGGTTT	AAGAATATTC	CTTTGTCTCA	AATGATAAGG	ACTTTATTGA
38751	GCTGGGTTTT	CTACTACATG	CCAATAGTTG	GTAGATCGCA	AGCTAAATTA
38801	AAAGTAACCA	AGAAGCAAAT	ATTTAAATTC	CATGTATAGG	AGCAAGTAAT
38851	CCTGACAAGT	AAACTCAGTA	AACCTAACAA	GAATTAGGTG	ATCCTGGTAG
38901	GAAGGGAGTT	TGAGGGAATG	TTACTAGTAA	TAATATTCTT	AAAGATTCCT
38951	AATCAGGCAA	AAGCAAAAAA	TCAAAATGAA	GTTCTCACAG	AAAAAAAAAA
39001	TTGATAGAGC	TTTATGCAGC	ATGAGTAAAT	CCCTCATTCC	TCGGGGGAAA
39051	TATCAAATAT	GATGAGATGA	TCATGGGAAA	AGAACTTCAG	CTTAGTTTTT
39101	AAGATATAAG	AGAAAGAGGA	TATTGATATG	TTTAATGATA	CAAAGACAGT
39151	TCCCAGGGGG	AAAAATTAAT	TTTAAAGCCT	TGTAGTACAA	AATAGATATT
39201	CACATAGACA	GATATGATTA	ATGGAAGGCA	ATAAATAGGG	GGAAAAAGAA
39251	AAGGTAATAG	GGCAATTTAA	AAGAAAAAAA	AGAGAGGTAG	ATATACAAAG
39301	AGACAAATTG	ATGAGTGAAA	AATGATTGAA	GTAGAAATAA	ATATATGGTC
39351	TATAAATAAC	TAGGTCATGA	AAGAAGACAC	TTGAGGATGG	TGATGCATTT
39401	AAACAACACA	AAAGTGATAA	TATGTAGACA	GAAATGAAGG	CTGAAAACAA
39451	TGGAGTTATT	TCAGAGCTAT	TTCCACAGCC	AGAAAAATA	CAAAATTCATA
39501	ATAAACATAA	AAACATAATA	CTAATAAAAC	TTGATGTGTC	AAATAAGACA
39551	AGAAAAATCA	GGGGGGCAAC	AATTCTTAAA	ATCTCTAAGA	AAAGGGCAAC
39601	ATTATTGTGA	GTGGAAGATT	TTTATTTAAC	TTTAGCAAAT	TTTAGGCAAG
39651	TTACAAAAAG	TAAAAATAGC	ACATGACTAA	TGTAATTAAT	ACATTAAGAT
39701	AATCAATGTA	TTAACGCAG	TTAACATTTT	ACAGAGAATG	TACACCCATT
39751	TTGATTAGCA	CATAGAATAT	TTACCCAAAA	TGACAGTATT	TCAGCATCCA
39801	AAACAGGCTA	TAAACTTAAG	CAGCAGATTT	TTATATGTGA	AAAATACAAT
39851	AAATCAAAGC	TCAAACCTTT	TAAAATTACA	AAAAAAAAAAC	CCCTCCTATG
39901	TCAACATTGT	GCCAGGTCTA	AATCCTACAC	GTATTACCTA	TTTATCATAT
39951	GTATTCTGTG	TCCTTCAATA	GAATATAAGC	TGCATAGATT	GTGGTCTTAC
40001	TTCACTACTA	CACCCCCAGC	ACCTAGGACA	GTGCCTGGCA	CATAGTGAGT
40051	ATTGTGGAAT	AAATAGATGA	TTGAATGTGT	GATGTGTTGC	TCATTTTATG
40101	ATGAATAAAT	AAGAAAATAC	TTTAATTAGG	ATGTCAACAT	TTTGCATGCA
40151	AATATGGCTT	CTAAAAATATA	TATTAATAT	ATTAATATT	GATCTTGCTT
40201	ATACTGTGAA	CTGTCTCAAA	AACATTTTCT	AAGTAATTTG	CAAAGTGCAG
40251	ATTTTATCTC	AGCTGTTATG	CAAATTACGT	ATTCTTAATT	AGTGACATAT
40301	TGGGAGATTT	TAATAAAGAA	AAATTCATTA	GTAAGCCTCA	TTCTTTTAAAG
40351	GAGAATGGTA	TCTTGGGAGG	TTTGTTGATA	AAAAAGATGA	ATACCTGAAC
40401	TACTTTGTTA	AACACTCACT	AAACAAGGTT	CTCACTCATG	GAGTTAGATC
40451	CACGCCCTTA	TCAAACCATG	ACAAAGATAT	TGTAAGTGGT	CTCTTAGTCA
40501	CCATTTTCT	TTCTTATATG	TGCATTTTAC	TTGCCCTCCAG	ACCAATATTA
40551	ACTAATGTAG	ATCTCTGTTA	AAGTCATTAT	GCTACTCAGT	AATCTTCATG
40601	GTTCCCTTTT	TCCTACTAAA	ATAAAATCCA	ATGTCATAAA	ACAGGTACAT
40651	GAGGCCCTTA	TAGTCTAGTT	TCAACCTGCT	TTCCCTATGT	TATTTTCTAA
40701	TATTTTCTTA	CCCTGCTCTC	CAATCCGATT	ATTATGCTTA	CTACCTCCTC
40751	TTACATTTTC	CACCTTCTTA	CCTTTCTTCA	TAACATTCCT	GACTCTGGAA
40801	GATTTGTFTA	GAGTTCATAT	CCAGGCTGGA	GCAGTTATCT	GATCAGCACA
40851	GAGAATGGTA	GTATTACTGT	TCCCTTTGAT	CCAGGCTCTA	AATTTTATT
40901	AAAGGAACTT	AAGGTTACTT	TTTTATACCC	ACATTGCTAT	ATGGGCTTAC
40951	ATTGAGTTTA	TATTCAACTA	AATACTAACA	GGTCTTATTT	ATATGACCTT
41001	CTGTCAAGCT	GAGTTCACAC	CATTCTTAGT	AAAGTTTGCA	ACCCTGTACA
41051	TTTGGCCCCCT	TAAATCACTG	CTTTCCCTTT	TGAAAAACAA	ATATCTCTTG
41101	ATTATACCAT	CTCCTCTCCA	TTTCTGCTAC	TGCCTTAGCC	CTTACCACCT
41151	TAAGCCCTTT	ATGAGACTAG	CAGAGAGAGA	GTAAAGAGAG	AAGCGAAAGA
41201	AAGAAGGAAG	AAGCATTGTT	CCTCACATGT	GGACTTATGT	TCAGTCCCCT
41251	CCTCCTTCCA	AACTATGTCC	TATATACTA	GCAAGGAAAA	AAATCATTTG
41301	AAAATTAAAT	CGAATGATCA	TGACTCCCCC	CAGACAAATT	CCCCTCAATG
41351	CCTCCTTGTT	GTCTTCACTG	TAGCCTCAGA	TCTGATGTAA	TTTATTTCCT
41401	ATCCTCATCT	CCCTCCTTAT	TTTTCATCTC	TTATCGTCAA	ACAGCTCACA
41451	CTCTCTGTCT	CTGGCCTTTG	TATTTGTATT	TCCTTGAGAT	GACAACATCA
41501	TTTCCCAGCT	TCTCTTCTCG	GCTTACTGTA	ATTCCTTCTT	CAAGACTCAG
41551	CTCTGGCACT	TCCTCCTCTA	CGAACTTTT	CTTGGCACCC	TATAGTAGAA

FIGURE 3, page 13 of 27

41601 TGTGCAGGTG CTCTTGCTCT CTGTTCCAGT GACACCAGAT TTACCTCTAT
41651 CATGATGCTC ATTATGCGGG TCTGAATTGC CTGCTCACTT TCTCTCTCCC
41701 CAGAATTAGA CTTTGAGCTT CTTGAGCTCC TTGAGACCAA TGAGTTTGTC
41751 TTTTCATCCCT GTAACCTCTAG AGTTGGAAGA GTGCCCAGGA GTTTGTCAGT
41801 TTATGGTGCC AGTAAACTA TTCCTGATTT TTCTCCTTGT TTATCCAAGA
41851 AGAGTAAAGG GCAAGATAAA AAAGGAATGT GATGGAATTC AATTTAAGCA
41901 AAATCAGGAT TTCAGCCTTT TGATATTTTA ACTAATTTAG TGAGCATTTA
41951 TATTTTGCTA TGCATTGTCA TTCCATTAGT ACAGGTGACT ATAATTAAAG
42001 CTTTCATGAG ATTATTTTGA TTCACCCTTA TCGTAAGACT AAAAATGAAA
42051 CAGACACAAA TAATCTGTCA TAAATGGTGA TTCTCTGGGA CCCAATTTTT
42101 TGGAGCCAGT AGTGAACAA GCATTGGATT TTCTGGGCTG GGAAACTGG
42151 AGATTTCAGT GTCCCTATTG ATTTGCCTTC TTTGGAAAAT GACTGGCTCA
42201 AAGACAACTG GGCCTTGTC CTCTATCATG GCCATCTTAA ATGTTATTTA
42251 ATACCAATAA TCAGTAATAG GTTTTACTGG AATGACGGAG TTGTGTAATC
42301 TCTGGAATTT TTCTGAAGAT TTCTAGTGCC TATTTCTGAT ATGGTTTAAG
42351 CATATATCTG GTCAAAGCTA GTCTCTCAAG GGTCCATCCA GTTAAGAATC
42401 TATCATCATT AAGCCTCAAA CATTCTTAAA ATAATGAAGG GTTCCTCTTT
42451 CCACAACCTC CTCTTTACTT TCCTGATCAG TAAATTGACC AGAAGAAATT
42501 AACCTACTTA CTACTAAGT TTTATTCTTT ATATCAGCAA GTATGTATAT
42551 GTGTGTGTTT TAACAAATCT AAAAGTAGAT TTCTTATAAA CAAGTGATC
42601 AGCTTTCCCT TATAGTACCT AGGTAATTAT CAATTGATTA ATCTGTATAT
42651 TTTAATGATT TGGCTCCTTC TCTAAAGAAG CAGAAAACCTA CTTCAAATC
42701 TAAGATAGCT GAGACTTCAT TACTTGTTGC AAAATAGAAT TTAAGTGGTA
42751 GAATCCCACT GGGGAGTACT AACATGAATA ATTACCATTA CAAACAATCT
42801 TCCAAAATGA ACAGTTTCAC TGCATTGATT GATAGTAGCA TCTTCAAATG
42851 TGATTTACAT TTATATCTCT AATGAAAATT AGTACGTACT TCACACTTTC
42901 TGATTTTCT ATGTCCCTTC TGTGGCAACA TAATGTCTTA TTTCTTCTAT
42951 TTGTATTTGT AAATTATAGA GTAATATTTG TGACAGGCAA TGGGTGAATA
43001 TGTTTTGCTA AGAGCCTACA CTTACATCAT CTGATTTTTT AAAATACCTA
43051 CTGCATTCCA CTCTACATTT CAATTTAATT TCCTTTAATT TGAAATGTGT
43101 CTTGAGTAAC TGCCATGGAT TTATCATAAT GCAATACTTT GTGTTTCCCA
43151 CTTTTAAAA TGTATTAAAA TTAGTGAAAA AAGTAACCTG GAGACAGCCT
43201 TGACTGAAAA AAACCTGAAT GACATTAAGT CAGAGTTACC ATATCTGGAA
43251 TATTTGTTCC ATGTTAGATG TAGCATGTGC TTTACATAAA TTATTTCCAA
43301 CTCTTGTAAT GAAGGAAGTA TTTTCTCTAT TTTGCTGGTG ACAAACCTAA
43351 AGCATAGAGT TAAGTAAATT GTTCAAGGGC CATGTTAGCA GGTGGCTAGA
43401 CCAATATTCA AATGGGGGTG GATCTGATGC CAAAGCCTGG GCCTTTATTC
43451 TAACACAAGG CCACAAGCCA CATTAACTTT TATTATTGCC ATTAATATGC
43501 CACAAGCTTA TATGTTACCT CTTACTGTCT AATCTTCCCA GACTCAAAAA
43551 AGACATAGGC TAAGACCAAG CCATATTAGT CTAGTTTTTC TGTCTAGTCC
43601 ATATCAGAAC ATATACTGTA AGTGCCCTAG TTCACAGGGT TAGGAATCAC
43651 TATATTATTT AGTTGGTAAT TTTCTTTTGT GTGGCTTCTG GCATAAGCTC
43701 TCTCTAGAAC CAGGGCCAAT TGTTTCTCTC TAATGACTTG GAGGGAGGCT
43751 AGCCTGAGGC TATCCTTAAA AGTGCAAGTT GATTTATCAT CTTTTCCTTT
43801 GTTCCATGGA TGAGATCCAA CATGCAGCTT CAACTAGCCT CACGGGGACA
43851 GATATGTTAA CTGATTTTCA TCCACAAGAA GAAACATTGG TAACAAGATT
43901 TGGCTATTTT CTAATGTTAT GAATGCAGTG TTTAAGCAAT TATTAAAGTA
43951 TATGCATACT TTTTAATCTC ATTCCCTGTG CCAAATATCT AGATAGATCG
44001 ATAGATACAT AGATAAATAG AAGGTGTAGT TACAATTGAA CATAGTCAAC
44051 AATAGAAATA GGATATGTTA AAGATGCTGA AAACCTCCAT AGCTTGAAAA
44101 GTTGTGAGAA TATGCAATTA ACAGTTTACA ACAGAAAATG GTTAACACAT
44151 CTCTTAACTA GGAATTAAAA CATTGAGAGT AAGACTAAGA GTCAAGCACC
44201 TGGCTAGAAT ATTAGAACCT GAGAGTGAAA TCTCATTTGC TTAGTGCAAT
44251 AGGACTTTAC TCCTATAATA GAGAATGAGT CCAGCTTATT AACATTGAA
44301 GAAATTATAG GCACTGTCTT TTTAAATAAA AATTGCAATT TATTTTATT
44351 AAGACAAGGA AGCAAAGCTG AACACTGCTT CCTATCTTTG GCCTCACTGC
44401 TTTTCTTACT TTTTGCCCTT GTCCTCTTTT CCCAGGTTTC TAGCCAATAC
44451 CACTTTTCAG GGCCTCAGTG GTTCCATTAG AGTAAAAGGT TCCACCATCG
44501 TCAGTTCAGA AAACAACCTT TTCATCTGGA ATCTTCAACA TGACCCCATG
44551 GGAAAGCCAA TGTGGACCCG CTTGGGCAGC TGGCAGGGGA GAAAGATTGT
44601 CATGGACTAT GGAATATGGC CAGAGCAGGC CCAGAGACAC AAAACCCACT
44651 TCCAACATCC AAGTAAGCTA CACTTGAGAG TGGTTACCCT GATTGAGCAT
44701 CCTTTTGTCT TCACAAGGGA GGTAGATGAT GAAGGCTTGT GCCCTGCTGG
44751 CCAACTCTGT CTAGACCCCA TGACTAATGA CTCTCCACA CTGGACAGCC

FIGURE 3, page 14 of 27

44801 TTTT TAGCAG CCTCCATAGC AGTAATGATA CAGTGCCCAT TAAATTCAAG
44851 AAGTGTCTGCT ATGGATATTG CATTGATCTG CTGGAAGA TAGCAGAAGA
44901 CATGAACCTT GACTTCGACC TCTATATTGT AGGGGATGGA AAGTATGGAG
44951 CCTGGA AAAA TGGGCACTGG ACTGGGCTAG TGGGTGATCT CCTGAGAGGG
45001 ACTGCCACA TGGCAGTCAC TTCCTTTAGC ATCAATACTG CACGGAGCCA
45051 GGTGATAGAT TTCACCAACC CTTTCTTCTC CACCAGCTTG GGCATCTTAG
45101 TGAGGACCCG AGATACAGCA GCTCCCATTG GAGCCTTCAT GTGGCCACTC
45151 CACTGGACAA TGTGGCTGGG GATTTTGTG GCTCTGCACA TCACTGCCGT
45201 CTTCTCACT CTGTATGAAT GGAAGAGTCC ATTTGGTTTG ACTCCCAAGG
45251 GGCGAAATAG AAGTAAAGTC TTCTCTTTT CTTCAGCCTT GAACATCTGT
45301 TATGCCCTCT TGTTTGGCAG AACAGTGGCC ATCAAACCTC CAAAATGTTG
45351 GACTGGAAG TTTCTAATGA ACCTTTGGG CATTTCCTGT ATGTTTGGCC
45401 TTTCCACATA CACGGCAAAAC TTGGCTGCTG TCATGGTAGG TGAGAAGATC
45451 TATGAAGAGC TTTCTGGAAT ACATGACCCC AAGGTAATAC TTCATTTTAC
45501 TTTAGCTTTC TTGATTGTCC ATTATAATTC CATATGTTGT ATCTTCTGCT
45551 GTAGTATGCT CATGTTCTTC CATCTAACAC AGGAATATTC TCTCAGCCAA
45601 GTATAGAGAC TAGTCCAAAA GTCTGTTGCC TGGTTTAACT AAATATTTCA
45651 TTGTTTGT TTGTTTGT CATAAATGAA AAAAAAGAC TGAGAAGTTT TGGGGAGTGT
45701 CTTTCTAGA GTAGTCTTT CTGATAGAAA TATCTATTAA TGCATCTTTT
45751 CTTGTATTA TTTGACCATC TGATAGCACA CCTATCAGGG AATGGTCTTA
45801 TAAGGTATTT TCACCCAAAG CACACCTTAA AACTGATGA ATTACTTATC
45851 TTGGGAATTA ATAAAAATAA ATTGGAAGAT CCATATTTTA AATAGCAAAG
45901 AATCTTTTTC ATCACTAAAA AGTGATACAA TGGAAAAGAT TAAATTTTAT
45951 TATAAGCACC AAAGTCAACT GCTAGGGAAC TCACTGAGTG TAGAACAAGG
46001 AGTATCAGAC TAACTGAGAT GGCAGAATTA GCTAAGGCCT ATAAAGTAAG
46051 GGGAGCTGCT CAGCTGACTA CCTTGCATAG AAGGGAGAGT GCCAGCAGTC
46101 CAAGGACATT CAAGAAGATT TTGTCTATCC AGGGTACCCT TGATATCCTA
46151 GACATCTGAC CCTAAGGGAA GAAGGAAGAG GAAGTGAGA GTGCAGGTAA
46201 ACAGCCAAAG CAGGTAATAC TTAGGTAAGG ACAGCCATTC CATGTTCTCT
46251 CTGGATTGAA CCAGGGCCCC TCTAAGTGAG CTGGGGTACA GAAAATTAGT
46301 CCAGCCCAAT AGGACTAGAG AGAGGGGACT GTCAAGGACC AAGGCAATTA
46351 GAACAGAGCT CAGGGGAGTA CTGCAGTCTT GATGGGAAAC AGAGTGCAGA
46401 TCTGAAGCTG CAGTGCAATC CAACATGTAG GATACATTAA GTAGAGATTG
46451 GAGAAAGGTT CAATTCAGCA GGCACACTCA GGACATACCA TGTCTAAAGC
46501 AACTTAAGCT AAGCTGAGCC TTTCATATTA TAAACATTC ACAGGCTTTT
46551 CCAATGCCCC CTTGTACAC CAAGTCTCAA TGTATTGATC TATTACTAT
46601 AAGTTACTAT TAAACATTTA AAATTAATTT CATAGACCAT CAACAAGTAG
46651 GACATTGTGA GCTATCTTTA CTAAATGATA GAATGCCCCA GAGGGCTGGT
46701 GGCAGCTTTA AAGATTTTTC ATAGATGGTT TCAATTGGAT GTAAGTTCTG
46751 TTTTGCAACC AAAAGAATGT AAGAAATTTG ACCCATATAT TGCAAACTT
46801 CTGATAAGTG ACATGAACCT CATGAGAGGA TTCAGCCAAC AATGCCTCAT
46851 TGACTAGGCA AGAAATTTTG TAACTTCTCA ATGAATACTC AGGGCTTTAT
46901 GTTAGGAGCT GGAATTCAGT GAACACAAAT AAAATCATTG GCATAAATAA
46951 ACGCATCACC CTAAAGGGAG ATGTTGGTGA TGCTTCTGCA TTCACATTCT
47001 GCACCTGGCAG CAGCAGCCTT TGTTTATTCT TTGCCCCAGG AGTCCTGTAA
47051 ATCTTCTGAA GGTTCAGC CTCACTAGAA ACTTAGATTA TTTGTGAGAA
47101 TCTCAACAAA GTGACTCCTA AATTATTAGC TCAAAATTA AAGTATTTAG
47151 TCTGATCTAG TAAAAAATAA AATCTAATAT ATGCCTGTTG TGGAGATTTT
47201 AGGCCATTAT CTTATGTAAA AAGATGAACA CAATACTAAC TAGAGCTTTA
47251 TTTATCAGAA CGAGTGATTG CCAAAATTA GCCAGGGATG CTATGCATGA
47301 AAAAGCTCTA AGAGTGATTA TGCTAAAGTA TAAAAATAA AATTATAAAG
47351 ACAAGCTATA CGCAGCAGTG AAATTATTTT TAAGCAAAAA GAAATAATCA
47401 TTTCTCTCT GACTCTACCA GAATAAAGAG TGAAATTTT TAATAAGTCT
47451 AACCAGTCC ACAACACAAA GCCAGCAGAA CTGGACAAAG TCAACTTTGC
47501 ATTACAACTA GTGAGTTCTT AATGAAATGG GACAAACCTA AATCTAACT
47551 ATTTTCAATT TGAGATAATA AATGAATTT CACTAAGTTT CCTAAAAATG
47601 TGCATGTGTG TACTAGCGTT GTTCTACGG TAAATATCT CTTTGTAGG
47651 TTAATCTCTT TCCATAAAAA GACACCAGA TTAATAGCT AAGCAATGCT
47701 TGACCCCGG ATTTAGAGG AATAATGTT GTGATAGAAT TTTTAGTTT
47751 TCATCCAAAT ATGTCATATT ATGTGGACTA CATAATGTCC TCACTCTCAA
47801 ACAGAGATAC AAGATAAAAA TTATTTCCAC ATCCCTCAGT CACTCAGAAT
47851 GTACTCTTAT TATACCGTAG TTAATCACAT GAGTGTTTTC TCTCTGTTA
47901 ATGCTAAATT CCTTGAGGAG GGCCTATACC ATTTCTTATA CTAAATAAGG
47951 TAGAGCAATT CAGAGGTAAA TGAGTCAAA TGTGTGACTT CAATCAATTA

FIGURE 3, page 15 of 27

48001	GGGAAGAAAA	GACATAAACA	AAAGGAGAAG	CAACTTAGAA	GGTTACAAAT
48051	ACGCCAAAGT	TTAGCAACAG	ATGAAGTACT	GTCCAAGTTC	AAGGAAAGAA
48101	TGATCACCTT	CAACTCATTA	CAATGGAAGA	AAGTTTAATT	AATAAGGTAA
48151	TATTTGTGCC	TATTCTTAAA	GAATAGTAAG	AATTCAGAAT	GTGTGTCCAG
48201	GGAGGAGAAA	TGGACATCCT	AGGAAAAGGG	AACCACAATG	GAATTGCTTG
48251	GGAGGTAATA	CTTTATATTA	CCTATAGAGA	ATACCGATTT	TGCTTTATCA
48301	CGAACTTCCT	ACTTGCTTTC	ACTTGCGACC	ATGCTTGTGA	GCTGCATGCT
48351	AATTCTCTGG	GTTGTTCTGG	GCTGATCTGC	CTCTTTTTTT	TCTGTAACCC
48401	CTGCAGAAGG	TAGGCCTTTG	ATTCCTAGTC	TCCTGAGAAA	ATAGAATTCA
48451	AGTATAAAGT	GGCTTTGTGG	AACTGCTGAT	TTTGAGTCCC	TGCTCTTTCA
48501	TTTTCACCTC	GTCTCTCCCA	AACCTTTCCC	CAACTCCTGG	GCCCCTTTTC
48551	AGGGCTTCC	TTAAACAGAA	TTTTTCTCCC	TATCCCTTTT	CCCTGGGAAC
48601	GATTCTCCCT	AGGCTTTAAA	AACAGTTGGG	CTCCTCTGCG	GAATTTACTA
48651	GTTGAACCTT	TTTGGACTCT	GCTCCTGTTT	CAAATTTGGG	GGTTTACCTA
48701	ATGTCAAACA	ATTAAAATGC	CCGTTGCTTA	AGCTGGGGCC	TTCTTTTGCA
48751	ACTACCCGGA	GGGAACCAAC	CCCTGTTTTT	TAAAGTGGCC	TAGGGGAAAC
48801	CTGGCCTTGG	GGGACCCCTG	GCCTTGGGGG	GNNNNNNNNN	NNNNNNNNNN
48851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NAACTGGCCA	AAGCCTTGGA
48901	AGGCCAGGGG	GAAGAGCAGG	ACGACAAGGG	TTAGGAAAAG	GAAGGCAGGA
48951	GCCAAACATG	TTGAACCCCT	AGAGGCCATG	TGAGGAGTTT	GGACTTCATC
49001	CCCAAGGGCC	CATGGGGGGA	GCCATTGATG	GGCATTAAGC	AAGTGAGGGA
49051	CATGATCAGA	TCTGAATTCT	AAAAATTCTC	TGAGTGGATC	AGAAAGGAAG
49101	AAAGGGAATA	TGTGGATGAC	AGAAATGATG	GTGACTCAAA	TAAGAGCAGT
49151	GGCAGTGTTT	GCAAAACATC	TCCATGTTAG	AGTGAGGTCC	CTGTCAGGCT
49201	AATAAGATGC	TTTCTCTCCT	GGAGTCTGCT	TATTTGCTAA	TCCTTGAAGG
49251	AACTTTCAAT	TCTCAGAACC	TCAGAACCCT	TGCCTTCACA	GAGCCAATAA
49301	GACAGATATG	GGAGCTGATG	AGCAAATACA	ACAGAAAGTC	ATGCAAAGGA
49351	AAAACACAAA	GAGACAGTCC	TCCAGCAGAA	TTCTGAGCCT	GGGTCTGCTG
49401	GTTCCCTGAT	TCCTTCCATG	CTTTGGAAAAG	GAACGATGTG	CTGGTGCAAT
49451	CAGAAAAGAA	AATGAGACCC	AGAAAAGCAG	GATTTCCCTA	AAAATCCTGG
49501	TCATAGACCA	TAAAAGTATT	ATCACAACAT	AATGTAAAGA	ATGTATTTCT
49551	TGAAGGAATT	CTGTTTCTAT	ATGAGAGAAA	GGTATTTTAA	TTAATGATTC
49601	TACAGCCCTT	TTACCTTAAA	GAGAGGGTTT	GTTTAGAAAAG	CTTGAGTTGT
49651	AAGCTATACT	GTTAGCTTAG	TTTGTGCTAA	GTGAGATGAC	AAAGCTGTCA
49701	CCTTCCTCTA	AGTTCAAAGA	ACAGTTATCA	ATTAATTACT	CATCAAGTAT
49751	TTTCTTGAGC	AACTACAAAA	TGAAAGGCAT	TAGAAATACA	AGGATAAATT
49801	AAAATTAACA	TTATTCTTGT	TTACATGGAG	CTTCCAGTTA	GTGAAGAAAC
49851	TGAGAGTCTA	GACGCTAATC	ACATGATTAC	ATAATTAAAT	TTAAAACCTGC
49901	AACTGTGATA	GATTCTACAA	AAGATAGGTA	TATGTACTGT	AAGGGCTTAC
49951	GAAAAAGAAT	CTGGCCTACA	CAATAAGTTA	GGGAAGTAAT	TACTGAACTG
50001	AGAGCAGAAG	AATGCATATA	AGTTAAATGG	ATAAAGCAGG	ATAGAAAAAT
50051	AGACTAGATA	AAATGAATTA	ACTTGACAAA	CAGATATAGA	CTACCTGCAA
50101	GGTGCCAAAG	GTCGTGATAG	GTAACGTGAT	TACACAAACA	AATGAAAGTT
50151	GGTCTGCAAG	TTACAGAAAA	TAGGATGGGG	TGAAGCTTGT	TAATTCCTAG
50201	TTGTGAGAAG	TTGGTGTTAC	TTTAACTATC	CAAGCCTCAG	TGTTTTAATC
50251	TGCAAAATAA	GAAAATAACA	TCTATCTTGT	AAGATTATTG	TGAAAATTTT
50301	AAAAGCCCTT	AGCTTACCGT	CTGGCATACA	AGTAGCAAAT	CCAGACAGAG
50351	GGAGTCAGTA	TGATATAGTA	TCAGAAAGCA	TGTTACACAG	TTGTCAAAAG
50401	CTGAATGTCT	TTGATCAGGC	TGAATTCCAA	TCCCAGGTAC	AATTACTAGC
50451	TCTGTGACCT	TGGACAAATG	ACACACTTTC	TAATCTTCAG	TTTCTTTATG
50501	TAAAAAGATG	ATAGACATAT	CTACCTCACA	TTTCTTGGTA	GGATCAAAGG
50551	AGATAATGCA	CGGAAAGGTC	AAAGATTACT	TAAATAAATA	TTTGCCCTACA
50601	TATTTTATAT	GTACAAGCTA	CTATATCAGC	ACAGATAAAT	AAGTAGTTAT
50651	GTTTCTCTAT	TGTTGTGTGT	TTGGGTAGGA	GGGTGAAGCT	AGACAAGACC
50701	AAAGGCTTCA	CATAGAAAAT	TTTCACCTAG	GTCTTTGCCT	ACATATTTTA
50751	TATGTACAAG	ATACTATATC	AGCACAGATA	AATAAGTAGT	TTTTTCATTG
50801	TTGTTGTTCT	GGGTAGGAGG	GTGAAGCTAC	ACAAGACAAA	AGACTTTACA
50851	TAGAAAAAAT	TGACCTAGGT	CTTGAAAAAT	TAATAGATGT	TGGACAGACA
50901	ATGCTGATAA	ATGTGCCTTG	CGCAAGCCAT	ACGACTGCAA	TTTGCTACTA
50951	TGAGAACAGC	CCCGGGGGTT	CAAGGCATAA	GGATAATATT	TTTAAAAGCC
51001	TGGTGAGTTC	CTCCTAATAA	ACATCATCGC	TTCAGTTGTT	GTCATGAGCT
51051	AGAATGCAAG	ATGATGTAGT	AGATAATAGC	ACATGCTTTG	AAATAAGATA
51101	GACATGGGTT	CTGAGCCAGT	CCCTACTACT	TACAACTGCA	TGACCTTATG
51151	CCCATTACTT	CACTTCTCTG	AGCTTCAGTT	TACTCATCCC	TAAGAGGAAG

FIGURE 3, page 16 of 27

51201 TAACAACAGT GCCCTCTTCA CGAAGCTATT ATGAGGATTC AGCACGATAA
51251 TGTATGTAAA GCTCTAAGTA TCTTGTGTTG CGCATAATAA GTGTTTAATA
51301 AATGTTAAGT ATTGTTATTA TTGTGATGT GGCATTAAGG TTATGCTGGC
51351 ATAAAAACATT AGAATTGTGTT CAGTGCATGG AACAAATTACA TTAAACTTAG
51401 AGCAAGCTAT ATTACTTACT TAACAGTGGA ATACCAAAGA AATACAATGA
51451 ACAACAGGTT TTAAGAGTTC CTATATGGCA GTGGTTGCAG GTATTTATCT
51501 TTGTCACCCCT AGTAACTTTG AGAACTCTAC AGAGTAGGCC TTCAATAAGT
51551 GTTGAATAAA TGAACGATTT TGCTGATTTT AAAATATTTT TTATACTTGT
51601 AAACATGGTA AGTGTTCCTG CAGATAATCT GTAAATAAAA AATACATCTG
51651 TAAATTC AACCAATGTGT TTTCTTCCAA GGTGTGTGTG CATGTGCGTA
51701 TGTGTGTGTG TGTATAATAT ATATGGGAAA TCATGGCATT TAAATAAATG
51751 TATACATGTT TTGTTTGTAT CATACCATAT GAAAGCTTTT TTCCACTTGC
51801 CATATGAGCA TTTGCCTTGT TTTATGAAAC ATACTTTACA AACATGATTT
51851 ATATGTTGCA TTACACTCTA TAAGAAATA CCATATAATT TATTTAACTA
51901 TTACCTTATT GTTACACCTT TACTGTGTGT CAAAACCTAAT TCATCTTCAA
51951 ATATATGTTA TTTGTACCA TATTCAGTGA GTTCTTAATC ATTTTATAGG
52001 TAAACAACCTT AAAAACCTAA TTATTAATAA AAATACTCTG TATTATCTCC
52051 CTCCTAGATG AGAACATTTT ATGAAAACAC TAAAAATAA TTCAATAAAC
52101 AAAATGTAA ATAGCCTAAA GGTGGCTAAA CACAAAAGTG ATGTAGTCAC
52151 AGCATTAGCG CAGGTTTATT GTAATCTCAG GATGTAAGGT TTTAACTTGG
52201 CCTTTATGTT ATTCTAACCA AGGAGTATCA TAATCTTTAT TATGAATGTA
52251 CACTTTGTCT AATATGCAGT TTACAATAAT GAGACTAATT CTACATGCCA
52301 ATTTGCATGG CTTCTAAAAG ATTCTAATAG GTTCCAATAT AAAGCAAAAA
52351 AAATGTTTT GTATTTTTTG TTTGTCGTTT CCATCTGTTT GCTTCTAAAG
52401 ATAGAGCAAT TTCTGATGTA AAAAGCATGT AGCCATGTCT GCACATTTCT
52451 ATGTACATGT TTCTGCCTGT GGGGGTTAGA AAAGTTCCGA ATTATTATTA
52501 AGTTTCAAAG AATTATGAAA AAAATGTTAA AAAACACTTC TAAGAATAAT
52551 TTTTATTTAA GCCATTCTT TTTCTCCCC ATAGTTACAT CATCCTTCCC
52601 AAGGATTC CGCTTTGGAAT GTCCGAGAAA GCAGTGCTGA AGATTATGTG
52651 AGACAAAGTT TCCCAGAGAT GCATGAATAT ATGAGAAGGT ACAATGTTCC
52701 AGCCACCCCT GATGGAGTGG AGTATCTGAA GTGAGTGCTA ACCTCTTGGA
52751 TCCAAAGAAA AATTCTCACT GAAGAGAAGT AATTAGCTA CTGACGCAGT
52801 ATATGTTAGT TTCTGAAAAT GACAGATGAA TATACCCACG TTGTGTTAAG
52851 TAATACTTTA CACTGTGTGA GATCCCAGGA GATGGTGTGA AATAGTGTGT
52901 ATTTAATTAT GTGACCTAGA TTACTTGTGA CCTGCATAGT CTCAAGTTGG
52951 TAGTAGCTTT GTTCAAAGAG TCAGTGGGGC CATTAGCATA GCAGATGGTG
53001 GAGGAAGTAA ATGTTGGCTT TACATCACTT GAGAATACAA TGGTGCTGAA
53051 ATAGTCAAAC ACACAGGGA TATGGCCAGA GGATAAAGTG TCCCAAGGCA
53101 ACGTTTTCTT TCCCAGTGAA TATATCTCTA AACCCATGGA ATGCCTCCTC
53151 TTGCCATGAG GAAATGGAGT TTATCTTAGA GATTTCCTGT AGAAAGGAAA
53201 TAAGATAGAA GACAATGATT CCCATGCCTG CATTCTTCTC ATCAGAATTT
53251 ATGAGAAGCA ATCATGAGAA ATCACACTGC CATGGCAGAT TATCAGAGCC
53301 TGTAATTCAA TGAAGTTGAA TACAAAGGCA GACAGTGCAG TGATGGGTCT
53351 GTTCTGTCTA GTCTTCCTGA GAAAAGGGAA AGAATGGTTC CTGAAAAACA
53401 GGAAGACATG AGGGTGAGTA GTCCTCTCCC TCCTGCTCGA TGGAAATCAAG
53451 ATAATAACAG ACATCCACAC CTCCAATTCC TAGAATTGTG CAGCATCAGG
53501 AAACCTGGTT CTCCATGGTC AGCATCAATA ATCTCCCCAA TGGACAGCAG
53551 GATCTGCCAC CTCAAATTCT TTTTAAAGAA AGAATAGAAA TAAATAAATA
53601 ATTTTATGGA ACATAAGGGT TTTGTCTTTC TCAACAACTT TAGAAACATG
53651 CCACTTAAAA AATTTTATGG ACTTTTAACT ATAGCTTAGA GAAAAAGCCT
53701 TGTCTCTCTA TATTTGCAAA ATTATACATG ATGTGTAAGT ATTATGAAAT
53751 GGCCTTTTCA ATTTTGCAAG AACATCAACA CATTACAGTC TCTCTCTGAC
53801 ATGAAGTTTA GAGTCCCTTT ACCTCCCAGA TCTTCTGTG TATTCTCTTC
53851 TTCAGGCGAA TTTATGGTTG AGAGAAAGAA TAAGATGTCA GGGTAGCAAT
53901 GGCTTCCAGC TCAATAGAAA TAGCAGACAA ACTAGGCTCT GCTGACAGTG
53951 TGAAAAGGGA TGAGATGAGC TACTGTGCA GTCCCCAGCA GTTCCACTCC
54001 ACTCAGGGCA TTCACGTATC TCAGGAGCTT TACCTGAGAA GGCCACGTG
54051 CCCAGCACTG GCCCTGCCCT AGCCTGAAGG GAAGCAATCT TCAGGAAAGC
54101 GGCCACAGAT GAAGGCCCAA GACAAGTCAA TTTTCCTTGG TAATAAACTA
54151 GCAAGTGGCA GAGTCAGGAC TAGGACCAGG TCTCTGGAAT CCAACTGCTG
54201 CTTCAGACTA GTCTGGGAAC GATGATGAAA GAGTAGGTCC TTGATGTTTG
54251 CAGAATAGTC CATGTTCCAG CAACATCTAT GTTGCAAGTTA GTATCTGAAA
54301 GCTAGTTAGA AATGCAGCAA CTCCAGCCTC ATCCCAAACT TACTGCATCA
54351 GAATCTTCAT TTTAACAAAG TCCCAGGCA ATTCACTGAT TGAGGTGAAA

FIGURE 3, page 17 of 27

54401 TTGGCATCTA GGCAGAGCTT ATCATTAATG CCCTCTCACC ACTTCTCTCT
54451 GGGCCTTAAC TCTCCACTTT CAGCCTAGTC ATTTCTGTG CCCTCAGCCA
54501 CCAC'TTCCCG CAACCACAGT CTTTCATGTTA CCTCCCTGGC ATCCCGAGAGC
54551 TCTGACCTAC AAGGCACAAC CCCTAGCATT GCCTGTGCAA GGAAC'TT'TC
54601 TACATATTGA ACCTGTCCTT TCCCTCTCCC ATCAAAATTC TCCTGGACTT
54651 AATTCTGCTC TCTCAGGGCC CTGCTTCTC ATTAACAGTT TTCCAAAAAA
54701 TTAAC'TCCTA CTAAAATGCT TATTCCTTTT ATTATGTTAA TATGTGACTG
54751 GTTTTCTTGT GATGTGTGAT ATGTATTTT AAATAATGCT TAAGAAAAGA
54801 CAGGGCATGA TTCTATAATA GAAATAACCA TTGGGGGCC TGTGAACCAC
54851 AACAGTGATT CAGCCAATCT AGAAGCTACT TGTAAC'TGGA TCCCACTTGG
54901 CCATTTCTCT ACCAGTGACT CAGGGTCCCA ATGGAGTCTG AGAGCTGACT
54951 GCTTTTCGCC TTTTCACGTA ACTGAAATTT ATCATAGCTA TCTGCACTTT
55001 GCAGTCTAAA ATCAAGAGTA GTTATTTAAG GAAGGATCCC AGAGACATTA
55051 GGCTTCATGA ATTACTGGTT TTAAAAAACT GAAATGAACC TCATCTTTTT
55101 TATTGTCATA TTGCTACCAC AAATATTTGT GGAATATTGG CAAGTGATAA
55151 CTTGTTGCTA CGTAGCTGTC AAGGTACATT ATGGTACTGT GGCAGTCGAA
55201 CTTTGATTGG AGAAACAGCT TTCAGCTCAA TTTTATTTT ATTGCCAGGA
55251 TTCCATTAAG ATTCCTTATC AACTTCTAGG AGACAATCCA CATCCCCAAC
55301 ACTTTCTAAA GCTTCCCAT ACTGTAGAGC TGGGAGATGC TTCATTTTGG
55351 TTAAAGTTAA ATTTGGGCCT CATTGTAACT TAAATCTGAT ACCCCTTTGA
55401 AAAGGGGATG CATTTTAAAT TGGTTATTTT ACTTATTTGA AGAGTAGGAT
55451 AAGAAAGCAA CGGTCA'TTGG TACCAAAAAG GGAAGCTGAC CTGCCAACTA
55501 TGTGTC'TATA CATGACCCAG ACAAAAGCCAT TCGTGTAAGG ATGTGTTTCC
55551 TGCCCTGATG AATCTTCTGG GTGTCTAGGG ATATCTTTCT CTTTTTGATT
55601 TTCTATGAAT TCTAGTCATA TTCTCTCTG TTTAGAAGCC AACTGTGTT
55651 AAATTAGAAC AGCCTCACCA CTGGATCTAA GAGAGGAAGG ACTGAGCCCA
55701 GAAGGGATAG AAAAGAGTTA TTCTTTTTCG AAAGCTGTTT GGACAACTCT
55751 AAGGGTAGAA AATCCTTTCT TTTTTTCAA ATTAATAAAT ATTTTATTT
55801 TTAAAAATA AATACCTACA CCTACACAAT AAAAAGGAAC TGAGGTAGTC
55851 ATCGCATGAG ATAGAAAGAA GTGTAATACA GAGTTC'TGGT TCCCAAGAAA
55901 CTTACACTTT AACTGGGGAG ATAATATAGT GCACAAAATG GTTGCTTTTC
55951 TGATCTTCA AGAATCCATT CCCTCTTTT GGTACAGCA TCCTGCATTT
56001 CTTCA'TCCCT CTTCTACTCT CACTGATTTA TATGAGGTGA ACCCCACCCC
56051 TGGCTCCAGG TGACACCACC TAGCCAATAA GAATGATAGT CCAGACTTTT
56101 TAATGATTGC CATTCTA'ACT GCTGTGAGAT GGTATCTCAT TGTGGTTTGT
56151 ATTTGCATTA CTAGTCCAAC CATTGTGGAA GTCAGTGTGG CCATTCCTCA
56201 GGGATCTAGA ACTAGAAATA CCATTTGACC CAGCCATCCC AT'TACTGGGT
56251 ATATACCCAA AGAACTATAA ATCATGCTGC TATAAAGACA CATGCACATG
56301 TATGTTTATT GTGGCACTAT TCACAATAGC AAAGACTTGG AACCAACCCA
56351 AATGTCCAAC AACGATAGAC TGGATTAAGA AAATGTGGCA CATATACACC
56401 ATGGAATACT ATGCAGCCAT AAAAAATGAT GAGTTCATGT CCTCTGTAGG
56451 GACATGGATG AAATTGGAAA TCATCAT'TCT CAGTAAACTA TCGCAAGGAC
56501 AAAAAACCAA ACACCACATC TTCTCACTCA TAGGTGGGAA CTGAACAATG
56551 AGAACACATG GACACAGGAA GGGGAACATC AACTCTGGG GACTGTTGTG
56601 GGGTGGGGG AGGGGGGAAG GATAGCATTG GGAGATATAC CTAATGCTAA
56651 ATGACGAGTT AATGGGTGCA GCACACCAGC ATGGCACATG TATACATATG
56701 TAACTAACCT GCACATTGTG CACATGTACC CTAAA'ACTTA AAGTATAATA
56751 ATAATAAAAT AAAAAAATAA AAAAAATTT TAAAAAGGA ATGATAGTCC
56801 ATTTCCATGG TAACAATATC CAGGGATGGG CTCAGGACAT AGTACAATA
56851 AAAGCAAATT AGACTAAGAG CTTTCTGAA ACTGTTCTCT ATAGAAAATC
56901 CTTTCTTAA TGGATGTATG TCTTTCACCT TTCCAAAAG AATTGGGAAG
56951 TGGCTGAAAA CAAAGAAATC GTTGCATGTA TTTTAGACAG TTATTTCTTT
57001 TTAAAACTTC TCCTTCCTTG CCCTCTTGT AGGTGGAAGC TCAGCCTATG
57051 CTGAGACTCA CCCTTCATCT GAACCTAGTC CCAACACTTA CTAGCTGTGT
57101 AACCTGGTTT AAGTTACTTC AATCCTCTGA GCCTCAATTT CCTCATCTGT
57151 TATATCACAG TCATTTCTGA GTGATAAAAG GTATAGAGAA CAATGAATGC
57201 AATGCCTAAC AACAAGAGT CCCTCTAACA GTGTAATAAG AATAAACGTT
57251 CTCTATGCGC TTCCTATTCA ATTCAGAGTG GCTCTGGCTT TACTGATGGA
57301 TTTAGAGTA ATTAAAGGAG CTGGTAGATA AACTCATTGG AAAGATGTCA
57351 TGCTGTCTTA TAAGAGTGCC TGTCTCCCT GGTCTGTAGT CTAGACATCA
57401 GTGAGAAGCC AAGACAGCTA AGTCAGCACC TAGGTAGCTT GTGCGGCCCT
57451 TAGTGTTCGG GTTCTGTCCC CTAAACAAAA GCCGCTGTC AGCCTTCATG
57501 CTTCTTCCC ATTAATGAAT CATT'TTCACT TTTCTCTCT GGTCTTAAAT
57551 ATAGGAATGA TCCAGAGAAA CTAGACGCCT TCATCATGGA CAAAGCCCTT

FIGURE 3, page 18 of 27

57601 CTGGATTATG AAGTGTCAAT AGATGCTGAC TGCAAACCTTC TCACTGTGGG
57651 GAAGCCATTT GCCATAGAAG GTATTAATCA GTCACCTCTG ATTCACTTTT
57701 ACTCAGGATG TGCTCAGTTT GCCAACCTAG AAAGTCACAA ATGCCAAAAGT
57751 CAGAAGCAAA GAGCTATTCA TCTTCCCTCG TTTTCATTTT CAACTCATAA
57801 GCACTTAGCT ATTAAGTTGC TGAAGTTAGG AATTTATTTT TCACCTATTTC
57851 AACAAATATT TACTTATCCA ATTTTTAGGG GGAAAAATCA TTGTTACCCA
57901 TATGATGTTG TTTCAGATAT CTGGGAGTGG TGGCACAGTG TAATAAATTT
57951 TAATTTAATC TGTATTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT
58001 GTGTGTTTAG TGGCAGGGTG TTGCTATGTG CTATGTGCCC CAAGCTTGTC
58051 TCCAACCTCT GGCCTCAAGT GATCCTCCTG CCTCAGCCTT CCAAATTCAT
58101 GGGAAATCTG TATTTTCTAA CAGCCAAATA CTCTAGCAAA TCTGACAGAA
58151 AAACCTAGCAT GATTACATTT TACAACCTGGG AGGGCCCCAA TATTGATCTA
58201 ATGCAATCTA CTCNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
58251 NNNNNNNNNN NNTGCCTGC CTTTTTCATTC ATGGCATATT ACTAAGGATC
58301 ATCCATATTG TTGCCTCTAG TTCTAGTTCA TATAGTTTTG ATTATGTAAT
58351 ATTCCATCTG TGCAAATAGT CCACAGTTTC ATTACTCATT CTCTGTCAA
58401 AGAACATTTG GGTGTTTCC AATTTTGTCT ATTATGAACA GTTCCTCTAA
58451 GAACATTTTGT GTACGTATCT CCTGATATGC TTGTGGCGGA GATCTTTGGG
58501 AATGGACTCA GGAACAGAAT TGCTGGTCTT AAAATATGTC AATATTC AAC
58551 CATATAAAAT AATACCAAAAC TGCTTTCCAA AATGTTTGTA TCAATTTTCA
58601 GTTTATATTA CAGCAAGCAA TATATAAATA ATCTTATTGA TCCTCATCTT
58651 CTTTAAACACT TAGTATTACT TCTTATTAT TTATTTTGG ACAATTTAAG
58701 TTGATTTTTA TAAGTCTCTG TTTATATTTT AGTTCCTATA CTGTATTGTT
58751 TATTCACATG GTTTTATTAT ATTATTCAAA AATTA AAAAT AAATTTAAAA
58801 AGTAAGAGAG GGTGATGCAT TAACACCGAT AAGAGAATGT CATCAACCAC
58851 AGACTAAGAT TAATCTGATT TTGTATATTT AAGGTTTCTA AGAGGGGTTT
58901 TGGAAGAGGT AGATAGGAAA TCCTAGCCCT GATAAAGACC TCAAAGATTG
58951 CCTCTAAGGA ATGTCTTAAT GGGAAAGGCA GAAGATCTTA AAATTTTCA
59001 CTAATGCACT GTGCACAGCC CATTCTCTC CTTTCCAAC TCAATTCATC
59051 TACTCAGAGA TGCAGCTGAT TTAAGGGTAA TCATGACTAG GAATGTCTTT
59101 GAGTGTCTTG AAAGAAAGTT GATGAAACT CATCACGCCC TTTTTTGGT
59151 CTGTTGGCAG TATCACACAA ATATGTACTG TGGTGGCAAT CTCTCAGGAA
59201 GGGTGTAAAA AACTCATCTG AGATTGTATT TTCTTCTAGG ATACGCATT
59251 GGCCTCCAC CCAACTCTCC ATTGACCGCC AACATATCCG AGCTAATCAG
59301 TCAATACAAG TCACATGGGT TTATGGATAT GCTCCATGAC AAGTGGTACA
59351 GGGTGGTTCC CTGTGGCAAG AGAAGTTTGG CTGTACGGA GGTATGGAAA
59401 GACTGTTGAA AATGGTGACA CGTTGTATAG CTGTACCTCA GAGAACATAA
59451 GGAAATGCTA TTACTTGTGC CTCATCATCT AGGTTATTGC ATTTACTAGA
59501 CTGTTGCATA ATATTTGGAT TATTTTITAC TTTGTCCAAA AAGCGTCCAT
59551 TCCTATAGGA ATTTACAGGG ATGTGGGTTT GTCTTAGATT TAAATGTGAT
59601 GCTATTTTGA TGAGTAAATA TCTAAATTTT TACTTTTCCC CATAACCTCT
59651 ATCCACAAGT GCAGAAGAAA TGCTGTCTG AATTACAGCA ATAGTATCTG
59701 AGATTGACAT GAACAGTGTT TGATTTAATC GTTAATTGAT GGACATGATG
59751 TGTACTTCAA AATATCCTGC AAAATCTAAT CAAAACATTG CTAACCTACC
59801 AGTGGTTACC ACCTAAAGAT AAAGCCTTTG TGGATGTAGA AAAATAGACA
59851 TACCTTAGCG AGTACTTCTT AGATTACAGA GTCATGGATT TTTGAAACTG
59901 AAAACAATCA ATTCCTTCT TTTATAGTTA GGAAAAGAAT ACCCAGAGAT
59951 AATCAGTGGT TTCCCTAAGG CACACAAATA ATCAGAATCT CTTTCTATTG
60001 CAGTACACTG ACTTTACAAT GTAATTAGAA AGAAACCTAT AAAATAAGGC
60051 AGAAATTGGA TGATTTAAAT TGGAGACTGG AAGGTATAAC CAGAGTCACA
60101 GGCTCAGATA TTTAAAAGAG GAGCCAAATA TACCTGAACA AGGTCAGCCT
60151 GACAATTTAA AAAAAAACA AACCACCTA AAGTGAATAA TAATGGTTTC
60201 AAGTAGTCTC TATTTGCCAA GAACCTCCAA AGTTGTATTT CAACACAGAC
60251 CTTTCTTCTG AGCTCCAGGC CATGTATCCA CTTTCTATT TGACATCTCC
60301 ACTTGAATAT TCTCCAGGCA TCTGAAATTT AAGTTGTCAA AAGCTGAACC
60351 TTGATCTTCC CTCCTAAATC TATTTTTCCT CCTGTGCCCT ACATCTGGT
60401 AATGGCTCCT CTGGCCATCT AGTTACTCAT TCTGGAAACT GGCAAGGATC
60451 CTTAATGCCT CCCACATCTC ATCTTCCACA TTCATCAACA AGCCATGTCT
60501 TTTCTTCTA TAAATATGTC CTCGAATATG CCCATTATT GTCATCTCCA
60551 CTGTTGTGAC TCGAGTCCAA GCCATCATTC TTCACCTGGA CCACTATAAA
60601 AGTCTCTTAA CTAGTTCTCT GCTATCACTC CTGCCTCTC AAATTTGTTT
60651 TCCATGTGGC AGATCATCTC TCTGCTTAGA AGCCTGCAGT GGCCTCCCAT
60701 GTTAATGGGA CAAAAATCTG AAACATTGGC ATAGCCACAA ATGCCCTGTG
60751 GGACCTGGCC CTGCCACCT CCCAGCCTCA TCCAGCCAC TTTCTCTTTG

FIGURE 3, page 19 of 27

60801 GCTTCCTGCA CAAGCTGGTG CCTCCTCTTC CTCAAAATC CATCCTCTGA
60851 CACTTGCTAA GTGAATCCTT ATATTATAAA CTTCCACTGA AAACCTCCCTT
60901 TTTTAAACCT GGTAATCCAA CCAAATGAAT ATATTCNNNN NNNNNNNNNN
60951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNGTGC AAAAATTCT
61001 GGCCTTCCCT GAGGACAGTC AGGGTGGGTC TCTGTCGATT TTTGTTAAGA
61051 ACCGTTCTAA TTA AAAAGAG TTCTAAGCAA TGCTCCATTA GTAATAAATA
61101 ATAGCAAGTC ACAAGGAACC CAGGTCTTAG TCGCATTAGA TTGCTAGTGG
61151 GCCTTGTTTC TCTCTCACTG CATATCTTTG TCCATGTGTG CCATTATACC
61201 AAAATACCCG AGACTGGGTA CTTTATAAAG AAGAGAAGTG TATTCCCAC
61251 AGTTCTGGAG GCTGGAAGTT CAAGATCAAA GCACCACCAA GTTTATTGTC
61301 TTAGTAAGGG CCCAGTCTCT GCTTCCAAGA TGGCATCTTG TTGCTGTGCC
61351 TCCACAGGGG ACAAATGATG TGTCTCCCA TGGTGAAAGG CTGGAAGGGC
61401 AACAAAAGGG ACTAGCTAGC TTCCTCAAGC CCTTTTATAA GGGCACTCAT
61451 CCCTTCATAA GGGCTCTTAA TACCAACCCT TGGAGTTTAG GTTTAAACAT
61501 ATGAGTTTGA GAGGGACATA TACATTCAA CTATGGCACT ATTCATGCCT
61551 TGTATACTTT TCCTTCAGCC ACAGCAGTCA GTTTCACAAC ACTATACCAC
61601 TGTCCAGGAA AGTCATGTCT TTTTTCCTTT GTTTATGCTA CCCTCAGGTT
61651 AGTATGTTCA CTTTCCCCAC CCAACCTCCC CATCCTGCCT GTCTTATTCA
61701 AATGTTATCT CATGAACCAG CTTCCCCACA AATCCCAGGA AGCATGGATC
61751 AAATGTATCA TGGAGGTATA TGTTTGTTTT CTCCCTAGAC TCAAGCTGTT
61801 TAAAAGGAAA GATATAACTT CTTTCTATCT CTATCCCATG ATACGCAGAA
61851 GAGTGCCTGG CACCTAGTAT TAACTCAATA TATTATGTAT TTTTACTTCT
61901 ACGGGATTGA CTCTTGTCAT TGATGTCAAT GGCTAAATAA TGGCTGTTTT
61951 AGAGTACATG TGAGGGTTTT CTTAATCTAC ACAAATCACT TAGCAAAGTG
62001 CTGGCCACAG ATATTTATTG TAACTACTAT TGTGCTATA ACTGCTGCTG
62051 TTTTGGTATA TAGCAGTAGA GCAGAAAGAG CCCAGGCTTT GGAGCTGAAC
62101 AGAAATGGAC TGAAAGACCA ACTTTGCCAC TTA AAAAGTTG CAGTAACTCC
62151 ACTTTCCTTA GTTATAAAAT AGACGTTATA TATCTATTTT AATGTAAAC
62201 ACTTGGCAGA AAGTAAGCAA CCAATAAATA TCATTAAAGT CCTTCTCTTT
62251 CTCTTCTATG CCTTTTTTGT GGCTTGCTC GGTTTACTAC ACTTCTCTGG
62301 TCTCACTTTC CTCCTCTGTA AATGAAGTAG ATGATTCCA TGTTCCTAC
62351 GAGATCAAAA TTCCAAAGTT ACTGAAGTCA CCATCTTTT CCCCTTAAAT
62401 CTACTCATTC TCTTCCCTGA GGGACGTTCT TTCCTTGACA GCTACCAGGT
62451 ATATTAAATT GTTTCAATTC TCTATCTCTA TCTCTCTCAA TCTCTAAGAG
62501 ACCATAAGGT GGTCCAGACC CAGGGCCTTG GCACAACTCC AGGGGCCATT
62551 TGCACAGTGG ATTAACATAC GAATAGCGCC TGTCAACTGA AGAATCATGA
62601 GCTTCATAAA TTTGGCCAGG AGATCTTTAG TTCTCATAAA GGGTTGCAGC
62651 CAGCAGGCCA GCCATCCTGC AGAATGGGAA GCATAGCCTC AGCAGAAGCT
62701 GAGAGCAAGC ACTTCAAGGG AGGGGTAAAA GGAACAGGA ATTTATGCTG
62751 AGTGGGGTGG CTGAGTATAC GTATTGAGTA AGCTATAGGA GGAGTCATAA
62801 ATATTTATGA AAAGAGATAC ATGCACATGT GCAGTTGAGC TTCATGCCTC
62851 TTCTTGGGCC CATGTTCAAA AAATGGTGGT GTTAGCATGA CCCGAGGGTG
62901 GAGATTTTGG TCTTCTGATG TCCAAATGTG AAGCAGAGGA CATGAAAACC
62951 CTCACATATG ATCCCCACA AGTTGGCCAA AACCATCTGG AGATTGTGGT
63001 CATTTTTAG GAAGGTGCA TTGGGAAACT GGTGAGCTGT CACACTGAAA
63051 CTGCAAAGAG GGAGGGAGAA TCTGGTTATG GCCTTAGATG ATTAGCTAAA
63101 GGTGATAAAG CAATGAGTTA TCGGTTTCTT GTTTTCCAGA GCTAGCTTTT
63151 GCTTACTTCT TAAGAATGAA TTATGGCTAA AGGTTAATAA GGAAGGGACA
63201 ACTGAGGCAT GACAGACCTC CCATCCAATC AAGGCCAGGA ACTCAGTTTT
63251 TAAGGTTTCT TCGGGGTCCC CATGGACAAG AGAAAGTTCTG TTCAGTCGGT
63301 TGGAGAGCTT TTAATTTTAT TTTTATTCT CACCCTCTAG GCATTGGGTG
63351 CTCCTCGGCA GCCCTCAGTC CAACCCTGGC TGAATTTCTT TCATGATGTA
63401 TAATGAAGGT CACAGAACAC ACAGGGAGAA ATAGTCTCCA GCTGTCCTTA
63451 AGTCCAGAAA AAATGAATAT CCATCTGAAA ACCAAAGAGT ACACAAGCAT
63501 TGGGCCAGGT ATATCATGTG ATTCATGGCC TTTTCTTTCC TATTCTGTCA
63551 CAGACTTTGC AAATGGGCAT CAAACACTTC TCTGGGCTCT TTGTGCTGCT
63601 GTGCATTGGA TTTGGTCTGT CCATTTTGAC CACCATTGGT GAGCACATAG
63651 TATACAGGCT GCTGCTACCA CGAATCAAAA ACAAATCCAA GCTGCAATAC
63701 TGGCTCCACA CCAGCCAGGT GAGTGCCACA GGTGTCTTGC TCCAATATTC
63751 TTAACTGTGA CAATTCCTAG GGATGGGAGG ATCCAGAGCC CTATGTCAGA
63801 CTACTAAGTT ATGTTACCAC AATAACGAGG GTGGGGTAGG CACTCTCATT
63851 AGAGCAGAAG GAATTCTCTC TATCCCATTA ATTCACTTTT CCCTTAAATT
63901 AAGGCATCCC ACAGTGCTCC TTCCTTTCCT ATAACCCTCC ACTGTTGTTC
63951 ATAATAGAAG AAATAGAAGT TTTTAAACCA GGTTTCACTT AATAGTGACA

FIGURE 3, page 20 of 27

64001	GTCTGAAAAG	ATGGTGTGAA	AGATTATTTT	CCTAGGTTCT	TTTCTTGCCA
64051	AAATTGAAAC	CTATTACGTG	AGTTTAAAGT	CCCTGGTCCC	CTTTAAGTGA
64101	CCCCACTTGC	CCCTGTTCTC	TCTGCTTTTT	CCCTGGTCCA	GAGGGGTTTG
64151	AGACCCATTG	GGGCCAAGGC	CAAGCTTCAC	AAGGCCGCAG	TTCTCTCCTG
64201	TTACCTAGAC	ATCAACGCAC	TGATTACAG	GGAGAACTGA	AACTGTCAAG
64251	CAGACTGGTG	ACTCAAGCAC	ACTCGTTTGC	ATATCTCTCC	CTGAATGAAA
64301	TTGGCATGGA	GACCCAAAAT	ATAGTCAGAA	GGCTTCATAA	TGATGGAAGA
64351	ACTCTAACAA	AGGGAGTGGT	TTCACTTGAC	CACAGCAAGG	CTGGATGAGC
64401	CATCCTGGCC	TCTCCTGTGG	TGAGACCACC	TCCTGCCCTC	CAGTGTACCA
64451	TGGATTACCT	CCCAGCAGGG	AGGCTGTAC	TTCTGCATTT	ACCTTATTCA
64501	TGTATTCATA	TACCCTCATA	CTAAACTCCT	CACCTTGGTA	ACACTTTTTT
64551	TTTTTCTAGA	TAGGTACTA	TCATTGGTTT	GGTTACAGTT	GAACCTGTGT
64601	TTTCCTGCTT	TTATCCTGGG	AAGTATTTGG	TCTTCTAGGA	GAACCTCACA
64651	GATCTTCCTT	ATTGTATGCT	AATATGATTA	ACTTAATTCT	CAAGAGTCCT
64701	CCTTCCCTTC	TTTGTTTTGG	TTGTCAATTA	ATTATATACT	TTTTACTCAT
64751	CTTTAGCAAT	TATGGAAGGC	CTTTTGCTAC	ACATTAGTTA	TGATGTGACC
64801	TCCTTATATC	ACATAACGTT	TTATAATACT	GAGCTTTCCA	TCAGAACACT
64851	TTTTTGTTC	TCCAGGTGGA	ATAATAACCT	CTAATTAGAA	TTTTTGTTC
64901	GTTGAGATTT	GAGGCTGAAT	GTTGATCTAT	TTTTAGGCAT	TTCTGGAATT
64951	TGTTTAAAGA	GGCCAAGGTT	TGTAACGACT	ACATATACTG	TTACATCTCA
65001	AGTGGTTCTT	TGTTCCCTGA	CCTATGTGTT	TCTCATAATT	TCTACCTGTG
65051	AATCTGTTTT	CTCCATCAGA	GATTACACAG	AGCAATAAAT	ACATCATTTA
65101	TAGAGGAAAA	GCAGCAGCAT	TTCAAGACCA	AACGTGTGGA	AAAGAGGTAA
65151	GAAGGGGCCA	ATGGCAACTG	TCTTTATATT	TGTAAAATAA	TCTTTAGAGA
65201	TCTAACTGTA	TAATTATTCA	GATCAAATCA	GGGCAATTTA	TCAAAAGAAT
65251	CAGTATAAAT	AGAGGGAAAT	AAAACATAAA	ATAAAAAATG	TATATGGACA
65301	CTAAAATGCA	GTGTACACAA	TATACTGTCA	TGGTTAGAAG	TGTAGTCATG
65351	GACTCTAGGA	TCAACCCTCT	GGATCCAAAT	CCTGTCTCCA	AGACTTATGA
65401	TATGTGTAC	TTTGTGTAAG	TCACTTATTT	AATCTCTCTG	CACCTCAAGT
65451	TCCTCATTTG	TAAAGTAGAG	ATAAAAAACAG	TACCTATTTT	CTAGAGTTGT
65501	TGTAAGATTT	AAATAAGATC	ATACATGTAC	ATCTCTGAAT	GAAAGGAAAT
65551	GCCTAATAAA	TATTAGCTAT	TATTATGTAC	AAAATACATG	TAAGAATTAA
65601	TGAATACCGC	AGGCAATTAA	TTCCATGTTT	TACTGTCTTT	TTGCATATT
65651	TCCACTCCCT	ACTCCTTTCT	AGCATTCCTA	GGAACAGAGT	ATTGGAAATA
65701	TGAAACAGAC	ATGTCATGCC	TAATTCATTT	CCTGGCACTT	TTCTACAAAC
65751	TCCCTAGCAA	AGAGCATCTT	ATTAATAGGA	AATAACAAAC	ATTAAATGCA
65801	TTAATGACAT	CTGAAAATCG	AAGCTCTTCA	CTCTCACCAC	ACCAGGCTGT
65851	GGATGACTGT	TCCCTATTTT	ATGGTGACTA	AAGATGTGAG	AAGCACTTGG
65901	GTCTGGTTCC	TGGCTAGTCT	CTGCTGCCTG	CTGCCTAAGG	CAACCTACT
65951	GATTCTTTTG	TACACCCAGA	GGCCTCAGAT	GAGGGCACAC	CTCTCATCAT
66001	AACAGAAGAA	AAAGGGATGG	AAAACAGGAT	TCTTTTGTGT	TGTATCTTTT
66051	CTGGGACTGC	TGCAGTCCCC	TTCTATGCAG	TCTCCATCTA	GCTTGTGGGA
66101	ATCATTTTCT	TTATCTCTTG	AAGTATCTCT	TTCCAGTCAA	TGAGCACTCT
66151	CCCCTCCCCT	CTCAGTCTGT	GGTATTCCTG	CATCATATTG	CAAGTGTGTT
66201	AGTGACAAGC	TGTATACTAG	TCCAGTCACA	GCTGTTCCTT	GACATGTTAC
66251	ATCTATTTTT	TCTATTTTTT	ACATAAAATTT	TTAATAACAG	TAACACAAGA
66301	AGACATAGCA	GCAAATGTAT	CATCTTACAA	TGAAAAAATA	TTTGTTTTCC
66351	AGCTATAATA	GAAACAGGAA	GCCCAATGAT	CCCATCTCCA	ACTGTGATAT
66401	GATTTCATATT	CACATCTTTC	TCACATAAAT	TGAAAACCAT	TTGTGTCTTT
66451	TGATGCAACT	TACCCAGTTT	TCTTGGCAGA	TTCCCTTCCT	GAACCCCTTA
66501	TTTTGAGGAT	CTAAGGAGAA	CAGGTGTTCA	TGGTTTAGCT	TGGGCTCACA
66551	TTTCCTGTGC	CTACCTCTAT	ACAACCCAAC	ATTAGCAACC	TGTCAAACAC
66601	AATGAGTGTT	TGGCGTACCA	TAGCCGTCAT	GTCTCTTTGG	AATAGTCCAG
66651	TGGAGTATTG	AACCTCAGTG	TTACATAAAT	GCTCCAGGGA	AGCCTATTTT
66701	ACCCATTTTT	AGTGTTAAAT	ACAGCTCACT	CACTGGTCAC	GTAACACTCT
66751	AAGACTGACG	AAGGCTTGAA	TCAAAGCAAA	GCCTAAATGT	TACTGAGGCT
66801	AGGAGTATAA	CACCAGCCTT	GGGTTATTTT	TTCCAAGTAG	AACTGAGTGC
66851	TTACACTCAG	CATTTGTCAC	CTTGCACTCA	TAGGTACCCA	CATCAAATAT
66901	CAGATGCTTG	GTGATACTAG	CAACTAGAAT	TTGGCACAAA	GTCCAGCATT
66951	TGTTTATTCT	TCTATATTAT	ATTACCAGAT	AGATATACAA	AGCTCTGGAG
67001	AAGACCAGTC	CAGCTATCTT	TACTTACCTT	ATCACTGTGG	CTGTCTAGAC
67051	AGTTGAAGAA	AATGTGTAGA	TGCTCTACTC	TCAGGTTTTT	CCTGCTATGA
67101	ACCATTGTAG	GGCATTAGAA	TGCTCTCCCT	CTCTTCTCTT	GGAAGTATAT
67151	CTATGCAAAT	GCTCATGCAT	GCTACAACCT	GACATCCCTC	CTCTGTGCCC

FIGURE 3, page 21 of 27

67201	CATATTTACT	GAACAAATAA	AAGAGCAAAT	AGATAAATGA	ATGAATTATT
67251	AACATGGGTT	TGAGGAAATG	CTTGAGAGAA	TTTGGGGCCA	TGATATGGAA
67301	GTAGGTATTG	TCCCTTTCTC	ATTTAATGCA	AAGAAAATAA	GGTACATATT
67351	GCAGGAGATG	ATTTATATAT	AGCCCTGGGT	TTATTCAACA	TGTGATTTCA
67401	CATAAGGTTT	TGGTCTATCT	TTCATCTCAC	TGGGTCCCCA	ATCAATACAT
67451	GTCACCCCTG	TTTTCCCTTT	CCTCTCACCC	CAAGACACAC	AAAAATTACA
67501	AACTACATAA	CAGCACAACC	AAGATTACTT	TAAGATTATT	CAAATTCAAT
67551	AGGAAAAGAT	TTGAAGAAAA	AAATTAAGG	GAATTATAAA	GCTAGAAGAA
67601	AAATTACATC	TCCTCTCTGA	CTCAGGTCTA	AAGCCAATGG	AGCTATAAGT
67651	GGGTTCATTA	CAGAACTTTT	ACCCAGCCCA	GGATACAAGA	AAACTGAGCT
67701	CTGGTACCCT	CTGCTCATTT	ATATAAACT	TAGACTATGA	GGCATGTTAA
67751	AGAACCACAG	GGTGGTTTGG	AGTGTGTGTT	TCAATGGCTT	GGGTCATGTA
67801	TAAGTTGGTC	TTTGCTATGT	GATAAATCAT	CCCAATACTT	AATGTTTTAA
67851	ACAACAGCAA	GTTGCTTACT	AATTCATGGA	TCAACTGTGC	AGTTTGCTGA
67901	TCTGAGCCAG	ACTTGGCTGA	ATTTGGCTGG	GTTTGCTCAT	ATGTCTGTAG
67951	CCAACTTTTG	GGGATGGGGA	AAGCAGCTAA	GGGCTGTCTG	GTCTATGATG
68001	GCCTCAGCTA	GTCAACTGGG	AAGCCTGAGG	ATTCTCTCCA	TATGGTCTAC
68051	CATCATGCAG	CAGGCCATTG	TGGACTTGTT	CACATTGCAG	CAGCAGGGTT
68101	CCAAGAGAGT	GGAAATGTGC	AAGACTTCTT	AAAAGTTTAG	GCTTGGAAC
68151	GCCACTCTGA	TATGTCTGCC	ACTTCTGTG	GCCCCAAACA	AGTTGTGAGA
68201	ACACTCCTAA	CTGAAGCGGG	GAGGAAAGCA	GATTCAGCAT	AGGTACAAGC
68251	TGCAAAGTCA	CATTACAGAG	GGCATAAATG	AAAGGAGAAA	AGAAGGTTTA
68301	TGGCCACTTT	TACATAAAGA	CTTTATTATT	CTTCTCTTTC	CCCTTCTCCT
68351	TCCAGATTGT	CCCCTTCTCC	TGGCAAGTAA	GAGTCCAGGA	AAAAAGTCAA
68401	TTCACTTACA	TGAATGGGAA	CAAAAACACA	ATGGCTTGGT	AGGGTGTTC
68451	TATTTAGTTT	TGTCTGTGG	TAGATTGCAA	AAGTTGTCTG	AACCCCTCCTA
68501	CTCCTCCTCC	TGCTTCTTCT	CCTATGGTGG	AAAAAATCAG	CAGCTATGCT
68551	GCTTAGGGCT	CTAATCTATT	CTTGTAGTGA	GAAACTCTCC	TTATCTCATA
68601	AGTATCAAAG	TGTATTTTCT	AAACAGGATC	AGCCTTCCCC	TGTGACTATT
68651	TGGCAATAAT	TCTCATGCTG	TCTATAGCCA	TCTCTCCATG	ATGGTAGTAG
68701	GTGATACGAT	GCAAGCCTAA	AACAGGATTG	CAAATTGCTT	TCTATATGAC
68751	TTTCATTATC	CCAGCAAGAA	ACTGAGGGCT	TTCTCGGGAT	TTTTTTAAGC
68801	ATCGGACCTG	ACCTGTCATT	CTCAACTCAC	ATAAAAATCA	TCCCTATAGT
68851	AAGAAACACT	TTGCTGAGAC	CTGTGGCTTA	TATGCTTTTT	TTCTCCCCAA
68901	GATCAAGTAG	TAAACATCAG	GATGGTCTTG	TGGGACTAAG	GATGAGCCAT
68951	GTTATGAGAT	CTGTGAGCAG	GTTGATGCTC	AGAACCCAAC	AAGTGAATAA
69001	ATAGATTTTG	CTTTTATTAA	AGCATCATCT	TTCAAATCAT	CAAACGTTTC
69051	AAGGTGTGGC	TAGTTTCTGA	GCTTCCCTTG	CAGAAAGGAA	ATTAAAAGCC
69101	ACCTGAGGTT	GTTTGCAAAA	AAAAAAAAAA	AAAAAAAAAA	TGCACCATAC
69151	CCCATCCTAT	CATCCCTTCA	AATGACACCC	AATTCCAGTT	TCAGAGCAGC
69201	ATGGGACTTG	AACTTTTGTA	TGTTTATGAC	TCTTTATTGC	CCCATGACAC
69251	CCTAGCAGGT	AGTCTGTCCA	TGGCTTTGTT	ACTTCATCTC	TAAATGCACA
69301	CCCAGCTCCA	TATTATTGCA	CAGGAAATGG	CTAACAGATG	AAGACAGCAC
69351	CTTGAGAGCT	GCAGAATGGA	AAGTAAATCT	AAAATTTCTC	TGTTTCCTAG
69401	GTTCTAATGT	GGACCCGTC	AGCTTACCGT	ATGGAATACT	TCCAATCTGA
69451	GTCATGACAA	CCGACGAAA	TACATCTTTA	GTGATGAGGA	AGGACAAAAC
69501	CAGCTGGGCA	TCCGGATCCA	CCAGGACATC	CCCTTCCCTC	CAAGGAGAAG
69551	AGAGCTCCCT	GCCTTGCGGA	CCACCAATGG	GAAAGCAGAC	TCCCTAAATG
69601	TATCTCGGAA	CTCAGTGATG	CAGGAACCTC	CAGAGCTCGA	GAAGCAGATT
69651	CAGGTGATCC	GTGAGGAGCT	GCAGCTGGCT	GTGAGCAGGA	AAACGGAGCT
69701	GGAGGAGTAT	CAAAGGACAA	GTGAGGACTG	TGAGTCCCTG	GTGACCACAC
69751	TGCTTCCCTT	TCTCAGTTCC	TGACCTTCC	CTGAGCCCTT	GAGACACTTT
69801	GTAATGCTCT	TTTGTAACCT	TCGACAAAGG	TGTGGGGAAG	CTGAGGTCTA
69851	GGTCTTCTTA	AAGGTCAAGT	CTGCTCTCCC	TCGCCTAAAG	TGCAGCAGCA
69901	GCTCCTCTCA	AGCTCACTCT	CTAGGTCTCC	AGGGTAGGAG	TGTTTTTCTA
69951	GCAAGAATCT	TAGTCAGGAG	TAAGCTCTGT	GCGAGAGATC	TGTGAATAAC
70001	CAGATAACCC	CAGCTGCCGT	TAACCTTTTC	ACCAGGTGCC	ACAGTAATAT
70051	TTCTGGTTT	TAGCCCTTTC	TCTGCACTAC	CAACAAGAGA	TAAAATTGTT
70101	ACTCACACTT	ATGTCTTACT	GGGTTGCTGG	TTTTCATCGT	AACACAGAAC
70151	GAGGTTATCT	AGGGTTGTAG	CTTTTGATAC	AACTCCCCGA	TCTAGATTTA
70201	TTCTTACATT	CTGAATGGGG	AGCAGGTAAG	AGCAGAGCAC	CTCCCACTGG
70251	GGGTGGGGTA	TTTAAAAATT	AACTCATTAG	TATCATAAAC	GTCAAGGATT
70301	GATTGGACCA	GGCAAGAGCC	ATGTTTTTGA	GAAGGTTCTG	GATCTCTGAC
70351	TCCATCCTGA	CTGTTTAGTA	AGAGCATGCT	TACACCCTAC	TGTGAAAAGG

FIGURE 3, page 22 of 27

70401	GGAGGGGATG	TGGTAAGCAG	AAACAGAAGA	CAGGCAGCAG	AGGCATTAAA
70451	AATGCATACC	ATGCTTTCAG	AACAAAAGCT	CTGGGCCAGA	AAGGCAATTT
70501	GGCTAAAAAA	TGAATAAGAC	TACTTCTAAT	GTAACATAAGC	ATCTCCACTA
70551	TGGTGTGTGC	CTTTTATAAA	GGAAAAGAGA	GAAAAAGGCA	AAGCAAGGTT
70601	GTGGCCTTAG	GTTTGACCTG	GAATATCCCT	TATTGCCTAT	AATGGAATAT
70651	GTGACACTGT	GGGTGAAATG	TTCTACACAC	CACACACTAG	GCCATTTTCA
70701	GATCAGCAGT	CACCCATCGC	TTAGCATAGA	AATCCCAAAA	CCTCCAGCCC
70751	GGGAACACTA	TAAGCTTCGA	CCATTACAGG	ATCTGCCCTG	CACTTTGCAT
70801	ATCTGTATAG	AAAATCAAGT	CAATCCCCCA	TCCTCACACC	CACTCATCTC
70851	TGAGGAGCTA	TGAACCTGGT	TTGGTCCCTC	TAATGATCCT	CCAGCCTCAT
70901	CTAATGCCCC	CCAAAGACTG	ATACAAGTAA	CCTCCCCTCT	GCTTAGGTGT
70951	CACTTTCTCA	GCTATCAAG	TTTAGGCAGC	AAGGGAAAGG	AATATGGGTC
71001	AGTTCTCAAA	TGTCAATGTA	GATAAGAGTC	ATCTAGTAGA	GAATCATCA
71051	GAGTGCGGAT	TGCCAAGACC	CTTCTCCAGA	GATTATGGGG	TTGGGGGTGG
71101	AGGTCTAGAG	GTGAGCTCAG	AAACCTACTG	TTAACCAACA	CCCCCAAGTG
71151	ACTGACACAG	GTGGTCTAAA	AATTACTTTT	CTAGAAACAC	CATTCTGGAA
71201	GTTTGGCTGC	CCACAGGCAG	GAGGAGAAGC	ATGAAGAGAA	AACCTGTTTG
71251	AGAAGTTTTG	TTTTGTTTTG	TTTTGCTTTT	TAATAATTTT	AGCACACATC
71301	TGCTGACTCT	CCTTCAACAT	CCTCACCCCC	ACCCCTGGGC	ACCATTTAGG
71351	ACAAGACTTC	CTTATTTATC	AATTACTTGA	TTTATCTTCT	CAGGACTCAT
71401	TGTTCCACCC	CCAACCAATT	TGAATGCCTA	CAATAAGTTC	AGGAGCTGTG
71451	CCAAGCACTT	TCCTCTTTTA	CAGCTGGAGA	TCCTGGAAA	GGTGTCTCAG
71501	TCACAAAAC	TCTCCCTCTA	CTACTGGATG	AAATGTCTGC	ATTTCCACCA
71551	AAATCTACCC	AGTCACCCAG	GGAATAACAA	CTTAAGCTGT	AGTTAGATAA
71601	CACCTAGTGA	TTAATTGGCT	GAGAAAACCC	TGGAGTGGAG	GGAGGCTCAG
71651	AGATACTGAT	ATGGATGTGG	GAGGGCTCTA	AAGTTAGAGG	TCACCAACTC
71701	CACAGATGAA	ACAGTTCAAT	AATGAGGAAA	CAGGTGAGCC	CTGAAAACAC
71751	AAAAGGACAG	TTCTGTGTTG	AAACACCCCA	TCCCCTCACG	TTCTACCCCC
71801	AGGCCCAGAA	GTAGGTTGCA	ACTGCCTTTG	GAAGATTTTG	CCCTTAGGCC
71851	ATCCCCACCC	ACTTGTAACA	GCTAAGAATG	CTGGAGACTC	TGCCACCATG
71901	CTCTGCGTGC	CCCTGAACCT	CTGTGCAGCC	CGGAAGGCTG	ATGTACAGGT
71951	GTACCTCAAT	CCACATTACA	GCCATGCTCC	TAATGTACAT	GGACATTTTT
72001	GTAACCTCAG	TCATATTCTG	ACTGTATTTG	AGAAGCTGGC	TGTTTAAGGG
72051	AACCCAGAAG	TGAATTCCTT	TGTAAAGTAA	AGCACCCCTT	TGTAATGCAA
72101	TTAATTATCC	CTTAATGTAT	CTGTTTGTGA	AGTCTGCATT	TTTGTATATC
72151	GGATTACCT	TAAGCTTCTC	TAGTGAGGCA	TTCTGAGCAG	TGGTGATCAC
72201	ATGCCAGATC	GCCCTGCCTA	TCCACAAAGT	AGATGACCAA	TGCACGCTCC
72251	TCAAACATCT	TTGGAGGAAC	TACCTGGCCA	AAACACTGGC	CAGGATGCAG
72301	CAAGCAGCAT	CAGGGGCTGA	CAGCAGGCTT	ACTGCCATCA	ACATTGCTTG
72351	AAATGCCTCT	ATGTTCTGAA	TAAAGAAAAA	CCATAATTGC	TTGTGGTGAA
72401	ACGAAGCAGT	CTTCATGTTA	AGTAGCAATG	GTTATTTTTA	TTGGTAGTAA
72451	CTGAACAGTG	TTTTGCAATT	TGTGAAACAG	TGTATTGTGT	TTTGTAATAA
72501	GATGTCATGA	AATGGTGGGT	CCTTGGAAAC	CTCCTTTCCG	TTCAGCTCTG
72551	CCCTGTGTTT	TTCAACTCCT	TTGAGGCTCA	AAAAAAACAC	AAAGATCAGA
72601	AGCTTTCAGA	TAGAGGGTGG	TATTCTGGTA	AAGAAGAAAG	AGATAAGGGA
72651	CGCTACCTTG	CTTTTCTGGC	ACAGGAAGCA	CATGATAAAG	CATGCTCAGA
72701	TGAGCTGGAA	CAGATATAGC	TACCTGGTTC	GTGTAAATAA	GAATAATCAA
72751	GGCCCAGAG	TGTGTATGCT	TCCAGGTGGA	GGAGAAAGGG	GAATCTCCCA
72801	AAATTTAAAA	ACAAATTGGA	AGAAATAACCA	GGACAGCCAA	GTGAAGCAGC
72851	CACAGGGACC	CAAGCAGTCG	AGGTCTTTAA	TGTGCCTGGA	GATGACTCTC
72901	TGCTATTCAT	GAATCTTGCT	ATTGCACAAA	CCCTATCAAG	AGCTGCTGCT
72951	TCCCTTCCAG	CCAGAAAAGT	GGTAAGCGGA	GCAAGTGCCA	AGCAGAACAG
73001	ACCTTATCAT	CTGGGTAACA	GACTTCTCAG	TGTTGGTGCT	GTGTCTGTTA
73051	GAGCCTTAGA	GCAAGTTAAG	CACTTCCTTG	GTGTGGGTAA	AGAATAAAGG
73101	GGAAAGAAAC	TACTTTAGAG	CCTCTTTTTC	TCCCAACTCA	TATTTTGTAT
73151	AGGAAAAACA	GAAAACCCAT	CCAGTTCTTC	AGAAATTGCT	TTCTAGGCAT
73201	TAATACTACT	TTACTATCTA	TACTGTTTAG	TTATTCCTTT	CTTACCCAC
73251	CTAACTATC	CATCTAATCC	AGGATTCCTT	CACTCTTTT	TTTAGTTAC
73301	TAATCATTTT	ATGAAAATAA	TGTATTTATA	AGTATTTTCT	TAAGGTTTGT
73351	GAAGAGTATT	TGCATTGTGT	CTTCATTTTA	ATGTGTTTGC	AATCGCTCCG
73401	CTCCAGGAAG	AACGGAATG	CTGTCTGTG	AGCATGAAGT	GAACGGGCTG
73451	TTTTGCTCCA	GCCACTTTTC	TTGTACAACC	ACATGGATGG	ATTAGATGTC
73501	CTCAGGTCTT	TTCCATCTTC	AGTTTCTATG	ACTGTGGAAT	AAATGTTTCAG
73551	ATAGAAACTT	CACTTTGGGA	TGTAAGTCTG	GCTTTGTCTT	TGGGGATTCA

FIGURE 3, page 23 of 27

```

73601 AATGTTGACA TGATACCACT TCCTTCTTAA TAGGAGACCC ATTAATGCTA
73651 TGATTTATGC TTATTTCCTT GCTATAGTCC AAAGAAGAAA CACAAGATAT
73701 GCTGAGAAAT CTCAGAGCTC AGGCATCATG AAGTAGAACT GAAATGGCTT
73751 CATCTGAGAT AGACATTCCA GGAAAAAGCA CAAGTTCAGA GGTCTCTAAA
73801 ATCCTGTACT GATCACCTC ATCAGTAATT CGACAAACAT TTGCTAAACA
73851 GCTTCCATGT ACGTGCCAAG TGCTGGAGAC ACAATAGTGA AGAAGATAGG
73901 TATGGTCCCT AACTTATGAC CTTTTTCTT TTTTTTTTTT TTTTTTTGAG
73951 ACGGAGTCTT GCTCTGTCAC CAGGATGGAG TGCAGTGGCA TGATCTCGGC
74001 TCACTGCAAC CTCTGCCTCC CAGGTCAAG TGATTCTCCT GCCTCAGCCT
74051 CCCGCCCGAG TAACTGGGAC TACAGGCGCC TGCCACCATG TCTGGCTAAT
74101 TTTTGTGATT TTAGTAGAGA TGGAGTTTCA ACATGTTGGC CAGGATGGTA
74151 TCGATCTCCT TACCTCGTGA TCCACCCACC TCGGCCTCCC AAAGTGTGGG
74201 GATTACAGGC ATGAGCCACC ACGCCACCC TCAATCTGAC CTTTTTACAA
74251 CCTATAAACA GGTAATACTG TAACAACATA CATATATTGG GTACTTATTA
74301 TAAACCATGA TCTCATTTAA TCTTAACAAC CCCACAAGAT AGGCACTATA
74351 GATGTAGTCT TAAGTAGGTA AATGAGACCT CCCAGTTTAC AGATAAAAAA
74401 ACAAGAGTCA GAGAACTAT GTAACCTGCC CAAGGTTGCA GAACTAGTAA
74451 TAGTAACAGA GATTTGTACA ACCATACAGG ATTCCGGTCA CTGCCTCACA
74501 ATTTTCTATT CTTCTTGAA TCCCCTTTTA GTCTTCTGC CTTACTGCTT
74551 CTTTCCCATG CCTCGGCTG GCCCTAGCT CCACAG (SEQ ID NO:3)

```

FEATURES:

```

Start:      1651
Exon:       1651   2349
Intron:     2350   29157
Exon:       29158  29762
Intron:     29763  44435
Exon:       44436  45483
Intron:     45484  52584
Exon:       52585  52730
Intron:     52731  57554
Exon:       57555  57670
Intron:     57671  59239
Exon:       59240  59391
Intron:     59392  63553
Exon:       63554  63718
Intron:     63719  65069
Exon:       65070  65149
Intron:     65150  69400
Exon:       69401  69740
Stop:       69738

```

MAP POSITION FOR ALLELIC VARIANTS:

STS that hits the 3' end of the cDNA:

WI-14669

dbSTS id: 37082, GenBank Accession: G23244

Organism: Homo sapiens Primer1: TCATAGAACTGAAGATGGAAAAGA (SEQ ID NO:5)

Primer2: GGAAGAACGGAATGCTGTCSTS (SEQ ID NO:6) location: 7633..7759 Chromosome: 9

GB4 Map: Chr.9

Reference interval: D9S176-D9S279 (104.9-120.4 cM)

Physical position: 326.27 cR3000 (P2.11)

RH details: RHdb RH62705

Typed by: Whitehead (see WI-14669)

RH mapping panel:

11000282865712	#	SHGCNAME	CHROM#	LOD_SCORE	DIST. (cRs)
	1	SHGC-9736	9	8.23	29
	2	SHGC-57676	9	6.4	40

blast match to HTG:

Coverage by HTG sequence AL137023.1:

Regions Covered:

1922 - 2971
2970 - 3118
3123 - 3234
3232 - 3385
3384 - 3550
3548 - 3629
3626 - 5481
5509 - 7789

HTG sequence, submitted 20-Jan-2000

emb|AL137023.1|AL137023 Homo sapiens chromosome 9 clone RP11-403A22 map q34.13-34.3, ***

SEQUENCING IN PROGRESS ***, 19 unordered pieces

Length = 184814

Score = 4389 bits (2214), Expect = 0.0

Identities = 2261/2281 (99%), Gaps = 1/2281 (0%)

Allelic Variants (SNPs):

3,248	g	a	atttgtgaacttaacggtgacaagtaataatgaggagatgaatctttaag [g/a]acaagac agagtccttatttagtaatgagttttctgccttttatatgtta (SEQ ID NO:7)
9,928	g	a	cttatgaaacaggagtgagcttattttggtgtggtagggtgagtagctg [g/a]aagagtt ccaaatctgaatcctcaaaacttggtgaatatgttattttttat (SEQ ID NO:8)
11,387	t	c	agtacaacctgcatgcaatctatgggtgtttttggacagaaggcctcaac [t/c]agaagcc aaacagaagtgtgttaataactccccagattaaaaagaaaagt (SEQ ID NO:9)
11,578	c	t	aaccagacattcttaaacagagattcctttaaacaaataatttgcttcta [c/t]atattgt aatgtaataatgggagcaaatatatacacagatccacacaca (SEQ ID NO:10)
11,731	a	g	cacattgtgttatacacataaagaaatgcttcaatgtgacctgaacatga [a/g]tgataaa tctagatccgaatttatctagtgtgccttcacctggccacaga (SEQ ID NO:11)
14,101	t	c	aggaattttctaacttgaaattgtggttatatctccaattctcaccttaag [t/c]taaaaa acttaagatgtcttgaaaaagtgtttttctcttacctataac (SEQ ID NO:12)
14,437	c	t	cttatcaaataataatgccttgagcttcatgccattcccttgctcaaaaac [c/t]atttac tataataataattccctttctttttccatgacccaacacttctg (SEQ ID NO:13)
16,732	T(17)	T(16)	ccattgaaacactgaaattttaaattggcctcctaaccatcctttaccacc [t/-]tttttttttttttttaagatggagctctcacactgttgccctgggctggag (SEQ ID NO:14)
18,612	a	c	ctcctgtcaaacaaagtatcgggaaatcagacaagagttcagatcttggt [a/c]agattag ccaagtctattcctaacttcctgttttactcactgctcatccg (SEQ ID NO:15)
18,968	a	g	acctgggaaaaaaaatcacatttggtagtttttaaagtatagaatttta [a/g]cctcact gaattccactatatttatatgctatgacctcatatatctgtttt (SEQ ID NO:16)
20,360	a	g	cttccaatttttgtttttctgggaggttattgttttctgttttatttgcc [a/g]ttgtaac tcaaggtctattacactgttttgctcatagtaatacactcaga (SEQ ID NO:17)
23,731	t	a	ctactcactgccttgctctacctcatttgttcttccacttagttctgtaac [t/a]ttgaagc agctctgaagtacagtgaacccatgacctggttgaagctag (SEQ ID NO:18)
26,282	a	t	ttaagccatcatgttgatagatcataaaatgacatctatcattctctgag [a/t]ctttcat aactgaaaaaggaataaatgcagtgtagagtcaggctagagt (SEQ ID NO:19)
29,047	t	g	tagacaggaagcatagttttccaaactatgggaattttatcccagaacta [t/g]gtatcac agtgaatttaaaggattaagcctcataagaaagcaaaagtacc (SEQ ID NO:20)
29,346	c	t	ctcctccttaccagaataattccaagttccaccttggttctatcatcaa [c/t]atcaccg ctaacctcccctccaccaggacctcttgagcttctacagat (SEQ ID NO:21)
29,542	a	g	ttggggtcatgccccctgaacttcggttggtgctgggagattcccagaat [a/g]tgaggag actgaggacagaggggtctgcccttaggactcattgctcatgga (SEQ ID NO:22)
29,577	a	g	ggagattcccagaatatggaggaaactgaggacagaggggtctgcccttagg [a/g]ctcattg ctcatggaaaaaacaacacagtctgtctttgagcactacgtaca (SEQ ID NO:23)

29,779	c	t	ctacaaatctcacttcaggacaatatttatcaaggtaggatgcaaggctc [c/t] gggtata tccccattcatagggccatgacagagagtaaaattcccctatc (SEQ ID NO:24)
32,135	c	t	aaaaaaaaattagctaagcatgctggcaccgactgtagtcccagctacg [c/t] gggaggc tgatgcaggagaatcgcttgaaacctgggaggcagagggtgcag (SEQ ID NO:25)
33,150	g	t	gctgacagaactgttaacatcttaaaatgttaatgaaatcaccaaaaaaca [g/t] ggcattt tcagctaggctttcagattagaaaagtcatttctcatggcaga (SEQ ID NO:26)
35,710	g	a	taagcattcttgtcctaaggacctctaccaacacaaaactggtaacccac [g/a] tatttca acatgtacttaaaagaaatgcagttgcattaaacatggaagcc (SEQ ID NO:27)
37,765	a	g	ctcaagaaatcacgcacatttaattgtttaatttggaactgtcccatat [a/g] tggagaa gaaatcaaacagaattggaccacaggtaagctctgtggctaa (SEQ ID NO:28)
38,468	g	a	ctggagggtctatggacattgtgaaatatcaaaattggctgtaagagtctc [g/a] aaagaca atccaaagagagaatagcttagggctcttgaaatgaaaaagagc (SEQ ID NO:29)
38,915	g	a	tcagtaaacctaacaagaattaggtgatcctggtaggaaggaggtttgag [g/a] gaatgtt actagtaataatattcttaaaagattcctaatacaggcaaaaagca (SEQ ID NO:30)
39,464	g	c	gtgataatatgtagacagaaatgaaggctgaaaacaatggagttattttca [g/c] agctatt tccacagccagaaaaaatacaaaattcataataaacataaaaaac (SEQ ID NO:31)
41,195	g	a	ccaccttaagccctttatgagactagcagagagagagtaaaagaggaagc [g/a] aaagaaa gaaggaagaagcattgttcctcacatgtggacttatgttcagt (SEQ ID NO:32)
44,478	t	c	tttcccagggtttctagccaataaccactttcagaggcctcagtggttccat [t/c] agagtaa aaggttcaccatcgctcagctcagaaaacaactttttcatctg (SEQ ID NO:33)
51,524	a	g	tatggcagtggttgaggtatttatctttgtcaccctagtaactttgaga [a/g] ctctaca gagtaggccttcaataagtgtgaaataatgaacgattttgct (SEQ ID NO:34)
54,016	t	g	tgagctactgctgcagtcacccagcagttccactccactcagggcattcac [t/g] tatctca ggagctttacctgagaaggccacgctgccagcactggccctg (SEQ ID NO:35)
54,405	a	c	cttcatttttaacaagctccccaggcaattcactgattgaggtgaaattgg [a/c] atctagg cagagcttatcattaatgccctctcaccacttctctctgggcc (SEQ ID NO:36)
55,007	c	t	cgctttttcacgtaactgaaatttatcatagctatctgcactttgcagtc [c/t] aaaatca agagtagttattttaaggaaggatcccagagacattaggcttca (SEQ ID NO:37)
55,156	t	g	tcatattgctaccacaaatatttgtggaatattggcaagtataacttgt [t/g] gctacgt agctgtcaaggtagcattatggtactgtggcagtcgaactttga (SEQ ID NO:38)
64,177	t	c	ttttccctggtccagaggggtttgagaccattggggccaaggccaagct [t/c] cacaagg ccgcagttctctcctgttacctagacatcaacgcactgattta (SEQ ID NO:39)
66,196	c	g	actctccctccctctcagctctgtggtattcctgcatcatattgcaagt [c/g] tgttagt gacaagctgtatactagtccagtcacagctgttccatgacatg (SEQ ID NO:40)
66,780	a	g	tcaactggtcacgtaacactctaagactgacgaaggcttgaatcaaagcaa [a/g] gcctaaa tgttactgaggctaggagtataacaccagccttgggttatttt (SEQ ID NO:41)
69,176	t	c	aaaaaaaaaaaaaatgcaccataccccatcctatcatcccttcaaatga [t/c] acccaat tccagtttcagagcagcatgggacttgaacttttgatgttca (SEQ ID NO:42)
70,027	a	g	catccccctccctccaaggagaagagagctccctgccttgcggaccacca [a/g] tgggaaa gcagactccctaataatgtatctcggaactcagtgatgcaggaa (SEQ ID NO:43)
70,419	a	g	taagagcatgcttacaccctactgtgaaaaggaggagggtgtggtgaagc [a/g] gaaacag aagacaggcagcagaggcattaaaaatgcataccatgctttca (SEQ ID NO:44)
71,332	c	t	aataatttttagcacacatctgctgactctccttcaacatcctcaccccca [c/t] ccctggg caccatttaggacaagacttccttatttatcaattacttgatt (SEQ ID NO:45)
72,153	a	g	aattatcccttaattgtatctgttttgtaagtctgcattttttgtatatcg [a/g] tttacct taagcttctctagttaggcattctgagcagtggtgatcacatg (SEQ ID NO:46)
72,711	t	g	cgctaccttgcttttctggcacaggaagcacatgataaagcatgctcaga [t/g] gagctgg aacagatatagctacctgggtcggtgaaataagaataatcaag (SEQ ID NO:47)
74,434	g	c	agtttacagataaaaaacaagagtcagagaaactatgtaacttgcctcaa [g/c] gttgcag aactagtaataagtaacagagatttgtacaaccatacaggattc (SEQ ID NO:48)

FIGURE 3, page 26 of 27

POSITION	Allele 1	Allele 2	Protein Position
3,248	g	a	Intron
9,928	g	a	Intron
11,387	t	c	Intron
11,578	c	t	Intron
11,731	a	g	Intron
14,101	t	c	Intron
14,437	c	t	Intron
16,732	T(17)	T(16)	Intron
18,612	a	c	Intron
18,968	a	g	Intron
20,360	a	g	Intron
23,731	t	a	Intron
26,282	a	t	Intron
29,047	t	g	Intron
29,346	c	t	Exon
29,542	a	g	Exon
29,577	a	g	Exon
29,779	c	t	Intron
32,135	c	t	Intron
33,150	g	t	Intron
35,710	g	a	Intron
37,765	a	g	Intron
38,468	g	a	Intron
38,915	g	a	Intron
39,464	g	c	Intron
41,195	g	a	Intron
44,478	t	c	Exon
51,524	a	g	Intron
54,016	t	g	Intron
54,405	a	c	Intron
55,007	c	t	Intron
55,156	t	g	Intron
64,177	t	c	Intron
66,196	c	g	Intron
66,780	a	g	Intron
69,176	t	c	Intron
70,027	a	g	Exon
70,419	a	g	Intron
71,332	c	t	Intron
72,153	a	g	Intron
72,711	t	g	Intron
74,434	g	c	Intron

296

N

N